UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE QUÍMICA E BIOQUÍMICA

Study on the regulation of the expression of alternative protein isoforms involved in carcinogenesis

Bruna Filipa Francisco Pereira

Mestrado em Bioquímica
Especialização em Bioquímica Médica

Dissertação orientada por:
Doutora Luísa Maria Ferreira Romão Loison
Doutor Marco Marques Candeias

2018

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE QUÍMICA E BIOQUÍMICA

Study on the regulation of the expression of alternative protein isoforms involved in carcinogenesis

Bruna Filipa Francisco Pereira

Mestrado em Bioquímica
Especialização em Bioquímica Médica

Dissertação orientada por:
Doutora Luísa Maria Ferreira Romão Loison
Doutor Marco Marques Candeias

This thesis was conducted at the mRNA Metabolism Group, Department of Human Genetics, Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA), IP, Lisbon, Portugal in partnership with the Molecular and RNA Cancer Unit (MaRCU),Center for Medical Education, Graduate School of Medicine, Kyoto University, Kyoto, Japan under the supervision of Doctor Marco Marques Candeias and Doctor Luísa Maria Ferreira Romão Loison, the internal designated supervisor in the scope of the Master in Biochemistry of the Faculty of Sciences of the University of Lisbon. This project was co-funded by FCT (PTDC/BIMONC/4890/2014 to MMC and UID/MULTI/04046/2013 to BioISI from FCT/MCTES/PIDDAC) and by INSA

"Science, my lad, is made up of mistakes, but they are mistakes which it is useful to make, because they lead little by little to the truth."

Jules Verne

Agradecimentos

Ainda me recordo do dia em que entrei pela primeira vez no Laboratório do Metabolismo do mRNA. Nervosa, ansiosa e desejosa de aprender. E sem dar por ela, cá estou eu! Um ano depois. Acabo este ciclo intelectualmente mais enriquecida, com um maior sentido de responsabilidade e maior capacidade crítica. Tal não seria possível sem a ajuda de todos os membros do grupo: Paulo Costa, Joana Silva, Rafael Fernandes, Patrícia Dias, Juliane Menezes, Manuel Carvalho, Neuza Bacalhau, Beatriz Damasceno e Rafaela Lacerda. Obrigada pelas sugestões, pelos ensinamentos e pela disponibilidade em ajudar a qualquer altura. Deixo ainda um especial obrigado à Rafaela! Sem ela, tudo teria sido menos fácil. Obrigada pela paciência e pela resposta às questões sobre ciência e biologia molecular. Obrigada pelos ensinamentos de bancada e correções deste manuscrito. Obrigada, também, pelo carinho!

Deixo ainda o meu agradecimento a todos os colaboradores do Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA) que de alguma forma contribuíram para esta tese. Destaco todas as pessoas do grupo de Oncobiologia, que sempre foram muito prestáveis na cedência de algum material ou reagente que era inesperadamente necessário.

Seria também impensável não agradecer aos meus orientadores, Doutora Luísa Romão e Doutor Marco Candeias. Sem eles, nada disto teria sido possível. Agradeço à Doutora Luísa por me ter recebido no seu laboratório, pelos conselhos e sabedoria. Agradeço ao Doutor Marco pelo seu otimismo, pela capacidade de ver sempre o lado positivo das coisas e, claro, por todo o acompanhamento científico. O meu mais sincero obrigado!

À Rute Brás, Francisca Aires e Natália Faria, obrigada pela vossa amizade e carinho. Se cheguei hoje até aqui, também o devo a vocês, que me ensinaram e ajudaram a crescer enquanto pessoa.

À Rita Francisco e Ana Francisco, as minhas irmãs emprestadas, obrigada por estarem sempre presentes, quer quando preciso que me chamem à razão, quer quando preciso que me animem. Obrigada ainda à Alice Santos, uma madrinha no verdadeiro sentido da palavra.

Por último, não posso deixar de agradecer aos meus pais. São eles que fizeram e fazem de mim o que sou hoje. Todos os valores e ensinamentos que me foram transmitindo ao longo da minha vida fazem-me querer dar sempre mais. São, sem dúvida, o meu orgulho. E espero ser, sem dúvida, o orgulho deles.

E claro, agradeço ao Vasco, por ter suportado todo o meu mau feitio após dias que correram menos bem. Obrigado pelas lágrimas que secaste, por me fazeres rir e, sobretudo, por me fazeres sorrir, todos os dias.

A todos, obrigada!

Resumo

Em eucariotas, a passagem da informação genética contida no ácido desoxirribonucleico (DNA, do inglês *deoxyribonucleic acid)* para proteínas funcionais envolve uma série de processos que estão física e funcionalmente interligados. Tudo se inicia no núcleo da célula, com a transcrição da informação armazenada no DNA para moléculas de ácido ribonucleico mensageiro prematuro (pré-mRNA, do inglês *premature ribonucleic acid*). À medida que as moléculas de pré-mRNA vão sendo formadas, estas vão sofrendo modificações que visam a sua estabilização e preparação para as fases seguintes da expressão génica. A estrutura *cap* (guanina metilada, m^7G) é adicionada na extremidade 5' do pré-mRNA, intrões (regiões não codificantes) são removidos, com consequente junção dos exões (regiões codificantes), num processo designado *splicing*, e a extremidade 3' do pré-mRNA sofre poliadenilação. Após este processamento, o mRNA maduro é transportado para o citoplasma, para que a tradução para proteína possa ocorrer nos ribossomas. De uma forma geral, esta inicia-se com a formação do complexo ternário, composto pelo fator de iniciação da tradução (eIF, do inglês *eukaryotic translation initiation factor*) 2 ligado a uma guanosina trifosfato (GTP, do inglês, *guanosine triphosphate*) e a uma molécula de RNA de transferência que transporta a metionina da primeira cadeia peptídica (Met–$tRNA_i$, do inglês *initiator transfer ribonucleic acid methionine complex*). Este complexo liga-se à subunidade ribossomal 40S e a outros eIF, incluindo o eIF4E, que reconhece a estrutura *cap*. A subunidade menor do ribossoma é, então, posicionada na extremidade 5' do mRNA para que possa fazer o rastreamento de toda a região não codificante da extremidade 5' (5' UTR, do inglês, *5' -end untranslated region*) até que um codão de iniciação em contexto favorável seja identificado. Aí iniciar-se-á a fase de alongamento da tradução com a síntese da cadeia peptídica, após o recrutamento e adição da subunidade ribossomal 60S ao complexo de tradução já existente. Quando o último codão é identificado, a tradução termina e a cadeia peptídica liberta-se do ribossoma. A par com isto, toda a maquinaria de tradução é reciclada, para assegurar novos ciclos de tradução proteica.

Em processos biológicos de elevado consumo energético, como a mitose e a diferenciação celular, e em condições de stresse, como a escassez de nutrientes e a hipóxia, esta tradução canónica de proteínas encontra-se comprometida. A redução global da síntese proteica é conseguida, normalmente, pela fosforilação da subunidade α do eIF2 ou pela sequestração do eIF4E, após um dado estímulo, com a consequente inibição da iniciação canónica da tradução. Neste sentido, as reservas energéticas da célula vão ser utilizadas apenas no processo biológico em questão ou na resolução de um dado stresse celular. Tal é conseguido pela expressão seletiva de determinados mRNAs. Estes transcritos, normalmente associados a funções de manutenção da homeostasia celular e a processos-chave da célula, conseguem ser traduzidos para proteína através de mecanismos alternativos de iniciação da tradução. Um desses mecanismos envolve locais de entrada internos do ribossoma (IRES, do inglês *internal ribosome entry sites*), onde, através de estruturas secundárias no mRNA, de alguns fatores canónicos da tradução e de outras proteínas auxiliares denominadas ITAF (do inglês *IRES trans-acting factor*), o ribossoma é recrutado para as imediações do codão de iniciação, sem o envolvimento da estrutura *cap* e sem ser necessário o rastreamento da 5' UTR do mRNA. Apesar das vantagens deste mecanismo alternativo da iniciação da tradução, uma vez que permite a recuperação da homeostasia celular através da gestão dos recursos energéticos da célula, a tradução mediada por IRES pode também ser nefasta. Por exemplo, as células tumorais aproveitam-se deste mecanismo para ultrapassar as condições adversas que se criam no microambiente tumoral (por exemplo, hipóxia, falta de nutrientes e stresse oxidativo) e proliferar. De facto, muitas das proteínas com expressão desregulada em cancro apresentam IRES no respetivo mRNA. Uma delas é o supressor tumoral p53. A presença de três promotores no gene *TP53* leva à expressão de três transcritos: dois longos, que permitem a expressão das isoformas FL-p53 (do inglês *full-length p53*), Δ40p53 e Δ160p53; e um mais curto, que permite a expressão das isoformas

Δ133p53 e Δ160p53. Apesar da existência de dois IRES capazes de regular a expressão das isoformas FL-p53 e Δ40p53 já ser conhecida, tendo já sido caracterizadas e definidas as suas estruturas secundárias, só recentemente foi identificado um IRES capaz de mediar a tradução não canónica do Δ160p53, uma isoforma que aparenta desempenhar funções pró-oncogénicas, na presença de mutações *missense* no gene *TP53*.

Considerando o exposto, esta tese teve como principal objetivo o estudo da regulação da expressão de isoformas alternativas de proteínas envolvidas no cancro, com especial foco na regulação da expressão da isoforma Δ160p53 através do mais recentemente descrito IRES. Na avaliação da capacidade de uma determinada sequência mediar tradução por IRES é frequente recorrer-se a constructos que contêm dois genes repórteres, designados bicistrónicos. O sistema bicistrónico usado nesta tese contem o gene da luciferase da medusa *Renilla reniformis* (RLuc, do inglês *Renilla luciferase*) e o gene da luciferase do pirilampo *Photynus pyralis* (FLuc, do inglês *firefly luciferase*). A sequência codificante da RLuc é o primeiro cistrão, sendo por isso traduzida de forma dependente da estrutura *cap* (tradução canónica), ao passo que a sequência codificante da FLuc forma o segundo cistrão. Entre os dois existe uma estrutura secundária designada *hairpin* (estrutura em grampo). Deste modo, só ocorrerá tradução da FLuc se a sequência clonada a montante desta conseguir recrutar o ribossoma de forma independente da estrutura *cap*.

O IRES capaz de mediar a tradução não canónica do Δ160p53 localiza-se nos primeiros 432 nucleótidos da região codificante desta isoforma e é inibido pela sua 5' UTR, ou seja, pela sequência codificante da isoforma Δ133p53 localizada a montante do codão de iniciação do Δ160p53. Tendo isto em mente, propusemo-nos avaliar o efeito das mutações *missense* mais comuns do *TP53* na capacidade de indução da atividade do IRES do Δ*160p53* na presença da sua 5' UTR. Recorrendo, então, ao sistema bicistrónico já descrito, foi possível observar o efeito inibidor da 5' UTR do Δ*160p53* na indução do IRES na linha celular cancerígena HeLa, uma vez que foi observado um aumento significativo da atividade da FLuc na ausência da 5' UTR do Δ*160p53*, quando comparado com a atividade desta na presença da 5' UTR. Mais ainda, de todas as mutações testadas (R175H, R248Q, R273H e R282W), apenas a mutação R175H conseguiu reverter de forma significativa parte do efeito inibitório da 5' UTR, em condições de stresse induzidas pela Thapsigargina, uma droga que induz stresse do retículo endoplasmático, com consequente fosforilação da subunidade α do eIF2 e inibição da tradução canónica. Tais resultados parecem indicar que as funções oncogénicas da mutação R175H vão além da alteração ou perda de função proteica, uma vez que esta parece também atuar ao nível do mRNA, induzindo a expressão do Δ160p53, uma isoforma que já mostrou ter importância na sobrevivência, proliferação e invasão de células cancerígenas. Continuando a caracterização da regulação da expressão do Δ160p53 pelo seu IRES, pretendíamos ainda identificar proteínas auxiliares da tradução alternativa desta isoforma, recorrendo a um sistema que toma partido das interações entre o RNA e a proteína da cápside do bacteriófago MS2. Clonando sequências de interesse do p53 a montante de repetições da sequência do *MS2* e realizando a co-transfeção dessas construções com uma outra que contém a sequência que codifica para a proteína da cápside, seguido de co-imunoprecipitação da cápside e dos mRNAs com as repetições do MS2, seria possível fazer a identificação de novas proteínas reguladoras da expressão das sequências clonadas através de espectrometria de massa. Nesta tese descrevemos várias estratégias de clonagem que foram desenvolvidas para clonar, ainda sem sucesso, as sequências de interesse do *p53* a montante de repetições da sequência do *MS2*, assim como possíveis soluções. Procedemos também a otimizações das condições de imunoprecipitação do Hdm2 (do inglês *murine double minute 2 human homolog*), uma proteína que interage com alguns mRNAs cuja expressão se encontra frequentemente alterada no cancro, como o *XIAP* (do inglês *X-linked inhibitor of apoptosis protein*) e o *p53*, regulando a sua tradução não canónica. Com esta técnica pretendemos, no futuro, identificar novos mRNAs regulados pelo Hdm2 através da sequenciação daqueles que co-imunoprecipitarem com esta proteína.

Neste sentido visamos identificar possíveis novos mRNAs detentores de IRES que também possam ter um papel preponderante no desenvolvimento tumoral.

Concluindo todas estas linhas de investigação, esperamos desvendar novos conhecimentos sobre a tradução mediada por IRES, assim como parte do papel deste mecanismo na carcinogénese. Provando-se a importância dos IRES no cancro, novas terapias direcionadas para estas estruturas secundárias do mRNA ou para as proteínas que auxiliam este mecanismo de tradução podem ser desenvolvidas.

Palavras-chave: mecanismos alternativos de iniciação da tradução, locais de entrada internos do ribossoma (IRES), Δ160p53, cancro

Summary

In eukaryotes, gene expression is a highly complex process composed of several steps. One of those is translation, the step that converts the genetic information contained in the messenger ribonucleic acid (mRNA) into functional proteins. Under normal conditions, most proteins are translated through the canonical translation initiation mechanism, which starts with cap structure recognition at the 5'-end of mRNAs, followed by 5' UTR (untranslated region) scanning until the appearance of an initiation codon in a favorable context. Yet, under unfavorable or energy-depriving conditions, such as endoplasmic reticulum (ER) stress, hypoxia, nutrient starvation, mitosis and cell differentiation, canonical translation is impaired and protein synthesis globally decreases. Nevertheless, some mRNAs, usually related to stress-responses, cell growth and cell death control, continue to be translated through alternative mechanisms. One of them involves internal ribosome entry sites (IRES), in which the ribosome is directly recruited to the vicinity of the initiation codon, without requiring the cap structure. This mechanism relies on mRNA secondary structures and can be assisted by some canonical factors and other auxiliary proteins named ITAFs (IRES *trans*-acting factors). Tumor cells take advantage of this mechanism to cope with the unfavorable conditions that characterize tumor microenvironment and proliferate. Indeed, many mRNAs containing IRES elements are found deregulated in cancer.

One example is the tumor suppressor p53. The *TP53* gene is the most commonly mutated gene in cancer and surprisingly, p53 mutations usually lead to the production of a mutant protein with oncogenic functions. The presence of three promoters on *TP53* gene leads to the expression of different transcripts expressing different alternative translation products: the full-length transcripts allow the expression of FL-p53, Δ40p53 and Δ160p53, and a shorter transcript produces Δ133p53 and also Δ160p53. While FL-p53 and Δ40p53 protein isoforms have been widely studied in terms of internal initiation mechanisms, the fact that Δ160p53 expression is mediated through an IRES element was not known until recently. Since this shorter p53 isoform was already associated with survival, proliferation and invasion of tumor cells, the recently identified Δ*160p53* IRES may have an important role in tumorigenic functions of Δ160p53.

Thus, considering the aforementioned data, we proposed to study the regulation of the expression of alternative protein isoforms involved in carcinogenesis, more specifically, the regulation of Δ160p53 expression through its IRES element, aiming to understand the role of IRES-mediated translation in cancer development. Knowing that Δ*160p53* IRES is located within the first 432 nucleotides of Δ160p53 coding sequence and that its activity is inhibited by Δ*160p53* 5' UTR, we evaluated the effect of hotspot p53 missense mutations (R175H, R248Q, R273H e R282W) in reverting the inhibitory effect of Δ*160p53* 5' UTR on Δ*160p53* IRES activity. To do that, we used a bicistronic system containing two reporter genes: *Renilla* Luciferase (RLuc) and firefly Luciferase (FLuc). RLuc is the first cistron and its expression is driven by cap-dependent mechanisms, while FLuc is the second cistron and is only translated if there is an upstream sequence capable of promoting its translation through cap-independent mechanisms. In our experiments, we were able to see the inhibitory effect of Δ*160p53* 5' UTR on Δ*160p53* IRES activity, in HeLa cells, corroborating the previous reported results. Moreover, from all tested p53 missense mutations (R175H, R248Q, R273H e R282W), only R175H was capable of reverting some of the 5' UTR inhibitory effect on Δ*160p53* IRES activity. This was observed for cells under 2 μM Thapsigargin-induced ER stress, which is known to impair cap-dependent translation. The obtained results seem to indicate that R175H oncogenic functions go beyond the alteration or loss of protein function, since this mutation also appears to act through an mRNA-dependent manner by inducing the expression of Δ160p53, an isoform that has already been shown to have importance in promoting tumorigenesis. Furthermore, in this thesis, we also aimed the identification of Δ*160p53* IRES auxiliary proteins, using a system that takes advantage of *MS2* RNA–

MS2 coat protein interaction. Cloning p53 sequences of interest, such as the Δ*160p53* IRES, upstream of *MS2* RNA repeats, followed by co-transfection of these constructs with that expressing the MS2 coat protein and co-immunoprecipitation, will allow the identification of *p53* mRNA-interacting proteins, through mass spectrometry. Here, we describe some of the cloning strategies used, though unsuccessfully, to attempt to clone p53 sequences of interest upstream *MS2* repeats as well as some possible solutions.

Moreover, knowing that Hdm2 (murine double minute 2 human homolog) interacts with several mRNAs commonly deregulated in cancer cells, such as *p53* and *XIAP* (X-linked inhibitor of apoptosis protein), regulating their non-canonical translation, we also performed Hdm2 immunoprecipitation optimizations, so that, in the future, co-immunoprecipitation of Hdm2-bound mRNAs can be performed. Then, new possible IRES-containing mRNAs regulated by Hdm2 that may also have a preponderant role in cancer progression will be identified by RNA sequencing.

At the end, concluding all these lines of research, we hope to unveil new insights regarding IRES-mediated translation of cancer-related mRNAs. In fact, understanding how IRES-containing mRNAs are regulated under different stress conditions and how the switch between cell homeostasis and cell neoplastic transformation is triggered, will provide important knowledge for the development of new therapeutic strategies.

Keywords: alternative mechanisms of translation initiation, internal ribosome entry sites, Δ160p53, cancer

Contents

Figure Index

Table Index

List of Abbreviations, Acronyms and Symbols

°C Degree Celsius (unit of temperature)
A Adenine
A-site Acceptor site
A549 Human lung carcinoma-derived cell line
ABC ATP binding cassette
ABCE1 ATP binding cassette subfamily E member 1
AEV Avian erythroblastosis virus
Akt Protein kinase B
Apaf-1 Apoptotic protease activating factor 1
ASK 1 Apoptosis signal-regulating kinase 1
ATM Ataxia telangiectasia mutated
ATP Adenosine triphosphate
Bag-1 Bcl-2-associated athanogene 1
Bcl-2 B-cell lymphoma 2
BiP Binding immunoglobulin protein
bp base pair (unit of length in Molecular Biology, 1 bp corresponds to 2 nucleobases bound to each other by hydrogen bounds)
C Cytosine (when referring to nucleotides)
C Cysteine (when referring to amino acids)
C-terminus Carboxyl terminus
C/EBP CCAAT-enhancer-binding protein
Ca^{2+} Calcium cation
Cat-1 Cationic amino acid transporter 1
CDK11 Cyclin-dependent kinase 11
cDNA complementary DNA
CHOP C/EBP-homologous protein
cIAP Cellular inhibitor of apoptosis protein
CITE Cap-independent translational enhancer
CMV Cytomegalovirus
CO_2 Carbon dioxide
Co-IP co-immunoprecipitation
Cyr61 Cysteine-rich angiogenic inducer 61
D Aspartic acid
DAP5 Death-associated protein 5
DBD DNA-binding domain
Dhx29 DEXH (aspartate–glutamate–X–histidine)-box helicase 29
DMEM Dulbecco's Modified Eagle's Medium
DMSO dimethyl sulfoxide
DNA Deoxyribonucleic acid
E Glutamic acid
E-site Exit site
E. Coli *Escherichia Coli*
eEF Eukaryotic elongation factor
EGFP Enhanced green fluorescent protein

eIF Eukaryotic initiation factor
ER Endoplasmic reticulum
eRF Eukaryotic release factor
F Phenylalanine
FBS Fetal Bovine Serum
FGF Fibroblastic growth factor
FLuc Firefly luciferase
FL-p53 Full-length p53 isoform
FT Flow-through
G Guanine (when referring to nucleotides)
G Glycine (when referring to amino acids)
g Earth's gravitational force
G1 First gap phase of the cell cycle
G2 Second gap phase of the cell cycle
GAPDH Glyceraldehyde 3-phosphate dehydrogenase
GCN2 General control nonderepressible 2
Gly Glycine
Gln Glutamine
GMP Guanosine monophosphate
GOF Gain-of-function
GTP Guanosine triphosphate
H Histidine
h Hour (unit of time, 1 h = 3600 s)
H1299 p53 negative human lung carcinoma-derived cell line
HCV Hepatitis C virus
Hdm2 Murine double minute 2 human homolog
Hdm2-A Alternatively spliced Hdm2 A isoform
Hdm2-B Alternatively spliced Hdm2 B isoform
Hdm2-C Alternatively spliced Hdm2 C isoform
HEK 293T Human embryonic kidney cells that expresse a mutant version of the SV40 large T antigen
HeLa Human cervical cancer-derived cell line
HRI Heme-regulated eIF2α kinase
HIF-1α Hypoxia inducible factor 1 alpha subunit
hnRNP Heterogeneous nuclear ribonucleoprotein
I Isoleucine
IgG Immunoglobulin G
IP Immunoprecipitation
IRES Internal ribosome entry site
ITAF IRES *trans*-acting factor
IQR Interquartile range
K Lysine
kb Kilobase (unit of length in Molecular Biology, 1 kb = 1 x 10^3 bp)
kDa Kilodalton (unit for unified atomic mass, 1 kDa = 1 x 10^3 Da =1 x 10^3 g.mol^{-1})
L Leucine
LAR Luciferase assay reagent
LB Luria-Bertani
M Methionine (when referring to amino acids)
M Mitotic phase of the cell cycle (when referring to cell cycle)

M Molar (unit of molar concentration, 1 M = 1 mol.dm^{-3})
m^{7}G 7-methylguanosine
Mdm2 Murine double minute 2
Met Methionine
Mg^{2+} Magnesium cation
MG132 Proteasome inhibitor
min Minute (unit of time, 1 min = 60 s)
miRNA microRNA
mL milliliter (unit of volume, 1 mL = 1 x 10^{-3} L)
MLH1 mutL Homolog 1
mM Millimolar (unit of molar concentration, 1 mM = 1 x 10^{-3} M)
mm Millimeter (unit of length, 1 mm = 1 x 10^{-3} m)
mRNA Messenger ribonucleic acid
mRNP Messenger ribonucleoprotein
mTOR Mammalian target of rapamycin
MW Molecular weight
N Nitrogen (when referring to atoms)
N Asparagine (when referring to amino acids)
N-terminus Amino-terminus
NaCl Sodium chloride
NaOH Sodium hydroxide
Neg Negative-regulation domain (when referring to p53 domains)
NES Nuclear export signal
ng Nanogram (unit of mass, 1 ng = 1 x 10^{-12} kg)
NLS Nuclear localization signaling domain
nm Nanometer (unit of length, 1 nm = 1 x 10^{-9} m)
NoLs Nucleolar localization signal
NT non-transfected
OD Oligomerization domain
ORF Open reading frame
P Proline
P-R_5'd160_F promoterless *Δ160p53* 5' UTR- and *Δ160p53* IRES-containing bicistronic plasmid
P-R_d160_F promoterless *Δ160p53* IRES-containing bicistronic plasmid
P-R_F promoterless bicistronic plasmid
P-R_MLH1_F promoterless *MLH1* 5' UTR-containing bicistronic plasmid
P-site Peptidyl site
p_12MS2 12 *MS2* repeats-containing plasmid
p5'd133p53_12MS2 Plasmid containing Δ133p53 coding sequence with part of *Δ133p53* 5' UTR and 12 *MS2* repeats
p5'd160IRES_12MS2 Plasmid containing *Δ160p53* IRES with *Δ160p53* 5' UTR and 12 *MS2* repeats
p5'd160p53_12MS2 Plasmid containing Δ160p53 coding sequence with its 5' UTR and 12 *MS2* repeats
p53 Tumor protein 53
p53α Full-length p53 isoform
p53β C-terminally truncated p53 beta isoform derived from alternative splicing of intron 9
p53Δex6 p53 delta exon 6 isoform derived from exon 6 skipping and appearance of a premature stop codon
p53γ C-terminally truncated p53 gamma isoform derived from alternative splicing of intron 9
p53ψ p53 psi isoform derived from alternative splicing of intron 6

PABP Poly(A)-binding protein
PBS Phosphate-buffered saline
PCBP Poly(rC)-binding protein
PCR Polymerase chain reaction
PCV Picornavirus
PDGF Platelet-derived growth factor
pDNA Plasmid DNA
PERK PKR-like endoplasmic reticulum kinase
pFLp53_12MS2 Plasmid containing full-length p53 coding sequence and 12 *MS2* repeats
pg Picogram (unit of mass, 1 ng = 1 x 10^{-15} kg)
Pi Inorganic phosphate
PIC Pre-initiation complex
PKR Protein kinase R
Pol Polymerase
Poly(A) Polyadenylated
PR_5'd160_F *Δ160p53* 5' UTR- and *Δ160p53* IRES-containing bicistronic plasmid
PR_5'd160R175H_F *Δ160p53* 5' UTR- and *Δ160p53* IRES-containing bicistronic plasmid, with p53 R175H mutation
PR_5'd160R248Q_F *Δ160p53* 5' UTR- and *Δ160p53* IRES-containing bicistronic plasmid, with p53 R248Q mutation
PR_5'd160R273H_F *Δ160p53* 5' UTR- and *Δ160p53* IRES-containing bicistronic plasmid, with p53 R273H mutation
PR_5'd160R282W_F *Δ160p53* 5' UTR- and *Δ160p53* IRES-containing bicistronic plasmid, with p53 R282W mutation
PR_d160_F *Δ160p53* IRES-containing bicistronic plasmid
PR_F Empty bicistronic plasmid
PR_HBB_F *HBB* 5' UTR-containing bicistronic plasmid
PR_MLH1_F *MLH1* 5' UTR-containing bicistronic plasmid
PR_MYC_F *c-myc* IRES-containing bicistronic plasmid
Pre-mRNA Messenger ribonucleic acid precursor
PSF PTB-associated splicing factor
PTB Polypyrimidine-tract-binding
PVDF Polyvinylidene difluoride
PXXP Proline-rich domain
pΔN6_HDM2 Hdm2-containing monocistronic plasmid
Q Glutamine
R Arginine
RLuc *Renilla* luciferase
RNA Ribonucleic acid
RNA pol RNA polymerase
rpm Rotations per minute
rRNA Ribosomal ribonucleic acid
RT reverse transcription
RT–PCR Reverse transcription–polymerase chain reaction
RT–qPCR Reverse transcription–real time quantitative polymerase chain reaction
RUNX1 Runt Related Transcription Factor 1
SDS Sodium dodecyl sulfate
SDS-PAGE Sodium dodecyl sulfate polyacrylamide gel electrophoresis

S Serine (when referring to amino acids)
S S phase of the cell cycle (when referring to cell cycle)
s second (unit of time)
Ser Serine
shRNA Short hairpin RNA
siRNA Small interfering RNA
snRNP Small nuclear ribonucleoprotein
SV40 Simian vacuolating virus 40
T Thymine (when referring nucleotides)
T Threonine (when referring amino acids)
TAD Transactivation domain
TAp63 Transactivation domain-containing p63 isoform
TATA box 8-bp AT-rich promoter sequence
TBS Tris-buffered saline
TCP80 Translational control protein 80
TISU Translation initiator of short 5' untranslated regions
TL Total lysate
TP53 Tumor suppressor p53 gene
Tris Tris(hydroxymethyl)aminomethane
Tris-HCl Tris-hydrochloride
tRNA Transfer ribonucleic acid
tRNAi Initiator transfer ribonucleic acid
U Uracil
Unr Upstream of N-ras
UTR Untranslated region
UV Ultraviolet light
V Valine
V Volt (derived unit for electrical potential)
v/v Volume/volume
VEGF Vascular endothelial growth factor
W Tryptophan
w/v Weight/volume
XIAP X-linked inhibitor of apoptosis protein
Y Tyrosine
Δ**40p53(α)** N-terminally truncated FLp53 delta40p53 isoform
Δ**40p53β** C-terminally truncated Δ40p53 beta isoform derived from alternative splicing of intron 9
Δ**40p53γ** C-terminally truncated Δ40p53 gamma isoform derived from alternative splicing of intron 9
Δ**122p53** Murine homolog of human Δ133p53
Δ**133p53(α)** N-terminally truncated FLp53 delta133p53 isoform
Δ**133p53β** C-terminally truncated Δ133p53 beta isoform derived from alternative splicing of intron 9
Δ**133p53γ** C-terminally truncated Δ133p53 gamma isoform derived from alternative splicing of intron 9
Δ**160p53(α)** N-terminally truncated FLp53 delta160p53 isoform
Δ**160p53β** C-terminally truncated Δ160p53 beta isoform derived from alternative splicing of intron 9
Δ**160p53γ** C-terminally truncated Δ160p53 gamma isoform derived from alternative splicing of intron 9
Δ**p53** Deltap53 isoform derived from alternative splicing of exon 7 and exon 9
µg Microgram (unit of mass, 1 µg = 1 x 10^{-9} kg)

µL Microliter (unit of volume, 1 µL = 1 x 10^{-6} L)
µM Micromolar (unit of molar concentration, 1 µM = 1 x 10^{-6} M)

1. Introduction

1.1 Eukaryotic gene expression

Gene expression is a highly complex process that converts the genetic information stored in deoxyribonucleic acid (DNA) into proteins, which are important for cell functions.[1] In eukaryotes, this process involves the following steps: DNA transcription to messenger ribonucleic acid precursor (pre-mRNA), pre-mRNA processing to mature messenger ribonucleic acid (mRNA), mRNA export from the nucleus to the cytoplasm and mRNA translation to proteins, which is followed by protein processing and folding into an active and functional conformation (Figure 1.1).[1] Although these events have been described as independent and sequential—in part due to their complexity, which requires the study of each step separately—, recent studies suggest that these events are more than linear.[1,2] A coupled network between them exists and not only each phase is physically and functionally connected to the next but also earlier and later phases of gene expression interact to each other.[1,2] Nevertheless, to facilitate the understanding of each event, they will be described individually.

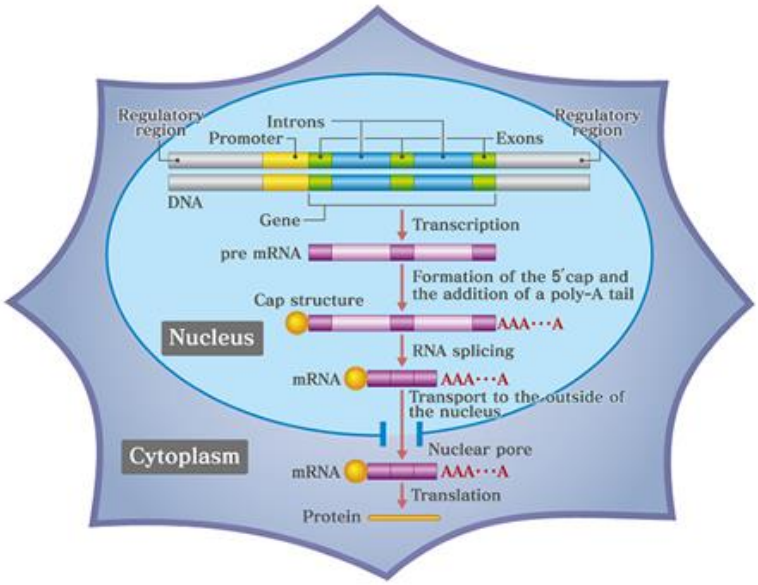


Figure 1. 1 — Schematic representation of eukaryotic gene expression. This complex process converts the genetic information stored in deoxyribonucleic acid (DNA) into proteins, which are important for cell functions. It involves the following physically and functionally connected steps: DNA transcription to messenger ribonucleic acid precursor (pre-mRNA), pre-mRNA processing to mature messenger ribonucleic acid (mRNA), mRNA export from the nucleus to the cytoplasm and mRNA translation to proteins, which is followed by protein processing and folding into an active and functional conformation. (Adapted from Division of Advanced Education in Sciences, The University of Tokyo.[3])

Transcription takes place in the nucleus of eukaryotic cells, and is the event in which the information encoded in the DNA is converted to RNA.[1,4] Although there are three types of RNA polymerases in eukaryotes (RNA Pol I, II and III), RNA Pol II, a 12-subunit enzyme, is the one responsible for transcribing protein-coding genes.[1,4] Transcription can be divided in three steps: initiation, elongation and termination.[5] During initiation, RNA Pol II and auxiliary proteins, known as general transcription factors, come together to form the preinitiation complex.[5] This complex recognizes DNA regulatory elements in the core promoter and, through the binding of some transcription factors to promoter consensus sequences, such as the TATA box, the RNA Pol II is placed near the transcription-start site.[5,6] Prior to RNA synthesis, DNA is unwound in order to allow the interaction between the single-stranded DNA template and the polymerase active site.[5] The next stage of the transcription cycle is elongation.[1,5] In this stage, elongation transcription factors are recruited to help the polymerase move

5' to 3' along the gene sequence and extend the nascent transcript.[1,5] In transcription termination, RNA Pol II reaches the end of the gene and transcription stops.[1,7,8] Here, the transcript and the RNA Pol II are released from the transcription site and from the DNA template, respectively.[1,7,8]

The nascent transcripts, called pre-mRNAs, must be processed into mature mRNAs before being exported to the cytoplasm.[9] Pre-mRNA processing comprises three phases—5'-end capping, splicing and 3'-end polyadenylation—that are tightly coupled to transcription.[9] In fact, transcription termination requires 3'-end processing signals.[9] The 5'-end capping is the first RNA processing event to occur on the nascent transcript, once 25–30 nucleotides have been synthetized.[7,9] First, the triphosphate on the first nucleotide is hydrolyzed to a diphosphate.[7,9] Then, a guanosine monophosphate (GMP) from a guanosine triphosphate (GTP) is transferred to the first nucleotide of the pre-mRNA, and, finally, the guanine from GMP suffers methylation at the N7 position.[7,9] The 5'-end capping confers stability to the transcript and is also important for mRNA recruitment to ribosomes, during translation initiation.[7,9] In splicing, a process characterized by two transesterification reactions, introns (noncoding regions on the pre-mRNA) are removed and neighboring coding regions (exons) are linked.[9] Consensus sequences on the pre-mRNA are required to intron identification.[7] They mark the beginning and the end of the intron (5´splice site and 3' splice site, respectively).[7] Also, around 100 nucleotides before the 3´splice site lies another important sequence, called the *branchpoint*, composed of a highly conserved adenosine followed by a pyrimidine-rich track.[7,9] Splicing is catalyzed by the spliceosome, which is constituted by small nuclear ribonucleoproteins (snRNPs).[7,9] The last pre-mRNA processing event is the 3'-end processing, which comprises mRNA 3'-end cleavage and polyadenylation.[9] Cleavage occur 10–30 nucleotides downstream the conserved sequence AAUAAA and around 10 nucleotides upstream a less conserved U- or GU-rich region and is followed by the addition of a poly(A) tail, which protects the mRNA from degradation.[8–10]

After mRNA processing, the mature mRNA must be transported to the cytoplasm in order to be translated.[1] Nucleus and cytoplasm are physically separated by the nuclear envelope.[1] So, to ensure communication between these two cellular compartments, the nuclear envelope has pores to permit the bidirectional transport of macromolecules.[1] For mRNA export, several factors bind to this nucleic acid, forming mRNA ribonucleoprotein particles (mRNPs).[1,11] Then, the proteins from mRNPs interact with proteins that line the nuclear pores and, along with other proteins from the export machinery, allow the transport of the mRNA to the cytoplasm.[11] After that, mRNPs suffer remodeling to prepare the mRNA for translation.[1,11]

1.1.1 Translation

mRNA translation is a very important step on gene expression.[12] This process takes place in ribosomes, which are large ribonucleoprotein assemblies that act together with several accessory factors to convert the genetic information contained in the mRNA into proteins.[13,14] Translational control is vital to regulate gene expression.[12,15] It enables rapid changes in cellular concentrations of the encoded proteins, thus controlling protein levels in time and space.[12,15] The steps of eukaryotic translation will be described below.

1.1.1.1 Translation initiation – the scanning model

Eukaryotic translation initiation is a very complex event that is assisted by more than 25 polypeptides.[12] This contrasts with translation elongation and termination, in which only a limited number of factors are required.[12]

Translation initiation of most mRNAs occurs via the canonical cap-dependent scanning mechanism.[16] According to it, translation initiation begins with the formation of the ternary complex,

which includes the eukaryotic initiation factor (eIF) 2 bound to GTP and the initiator methionyl transfer RNA (Met–$tRNA_i$).[17–19] eIF2 is a heterotrimeric factor composed of α and β subunits bound to a central γ subunit.[18,19] The eIF2γ subunit binds both GTP and Met–$tRNA_i$, while eIF2α and eIF2β are responsible for stabilizing transfer RNA (tRNA) binding.[18,19] Once ternary complex is assembled, it must bind to the 40S ribosomal subunit.[19] This process is aided by eIFs 1, 1A, 3 and 5, which altogether form a multifactor complex that stabilizes ternary complex and 40S subunit interaction.[19] Regarding eIF1 and eIF1A, they bind to the small ribosomal subunit, which is followed by 40S subunit conformational changes that will allow the accommodation of the ternary complex.[19] Concerning eIF3, which is nearly as large as the 40S subunit, it interacts with the small ribosomal subunit and with every component of the multifactor complex, functioning as a bridge between them.[19] At last, eIF5 serves as an adaptor between 40S subunit-bound eIF3 and ternary complex, through interaction with eIF2β and eIF2γ subunits.[19,20] The complex formed by ternary complex and 40S ribosomal subunit binding, together with the mentioned eIFs is called "43S preinitiation complex" (PIC) (Figure 1.2A).[19,20]

The following event is the recruitment of the 43S complex to the 5'-end of activated mRNAs.[19] mRNA activation is accomplished by the poly(A)-binding protein (PABP) and the eIF4F complex, which is composed of eIF4E, a protein that recognizes the 5'-end cap structure, eIF4G, an adaptor molecule, and eIF4A, an enzyme with ATPase and helicase activity. (Figure 1.2B).[16] First, eIF4E binds to the cap structure, then, eIF4A unwinds RNA secondary structures in the cap-proximal region of the transcript.[16,18,19] eIF4G and another initiation factor (eIF4B) are known to enhance eIF4A helicase activity.[19] They directly interact with eIF4A, with a consequent induction of eIF4A conformational changes into an active state.[19,21] Moreover, these eIFs also stimulate eIF4A ATPase activity, which is required for eIF4A helicase function.[19,21] eIF4G can also bring together the 5' and 3' -ends of the mRNA through interaction with eIF4E and PABP.[16,18,19] This may permit the coupling of translation termination and recycling events with subsequent rounds of initiation on the same mRNA.[19] After mRNA activation, PIC and mRNA attachment is mediated by eIF4G and eIF3 interactions and leads to the formation of the 48S complex (Figure 1.2C).[16,19] Once assembled to the mRNA, the 48S complex must scan the 5' untranslated region (UTR) base by base until it finds the first start codon in a favorable Kozak consensus sequence (Figure 1.2D).[16,18,19] This consensus sequence is characterized by a purine residue at position -3 and a guanosine at position +4, relatively to the +1 adenosine of the AUG start codon.[16,18] The 5' UTR scanning is dependent on adenosine triphosphate (ATP) and is assisted by eIF4A and other helicases, such as the DEXH (aspartate–glutamate–X[i]–histidine)-box helicase 29 (Dhx29), which unwind RNA secondary structures.[16,18,22]

Complementation with the anticodon of the Met–tRNAi results in scanning arrest and release of eIF1, eIF2–GDP, phosphate (Pi) and eIF5 from the PIC (Figure 1.2E).[16,19] The next event is the assembly of the 60S ribosomal subunit, which is accomplished by the GTPase eIF5B and eIF1A (Figure 1.2F).[16,19] After subunit joining, eIF5B hydrolyzes GTP, thus losing affinity for the initiation complex.[16] After eIF5B dissociation, the eIF1A follows the same path and the 80S complex is, then, formed (Figure 1.2G).[16,19] The eIF3 remains bound to the 40S subunit during the first steps of elongation and plays a role in downstream reinitiation events.[19]

In order to allow the formation of a new ternary complex, thus ensuring another translation cycle, eIF2–GDP must be recycled to eIF2–GTP.[16] This is accomplished by the guanosine nucleotide exchange factor eIF2B (Figure 1.2H).[16,18,19]

[i] X denotes any amino acid.

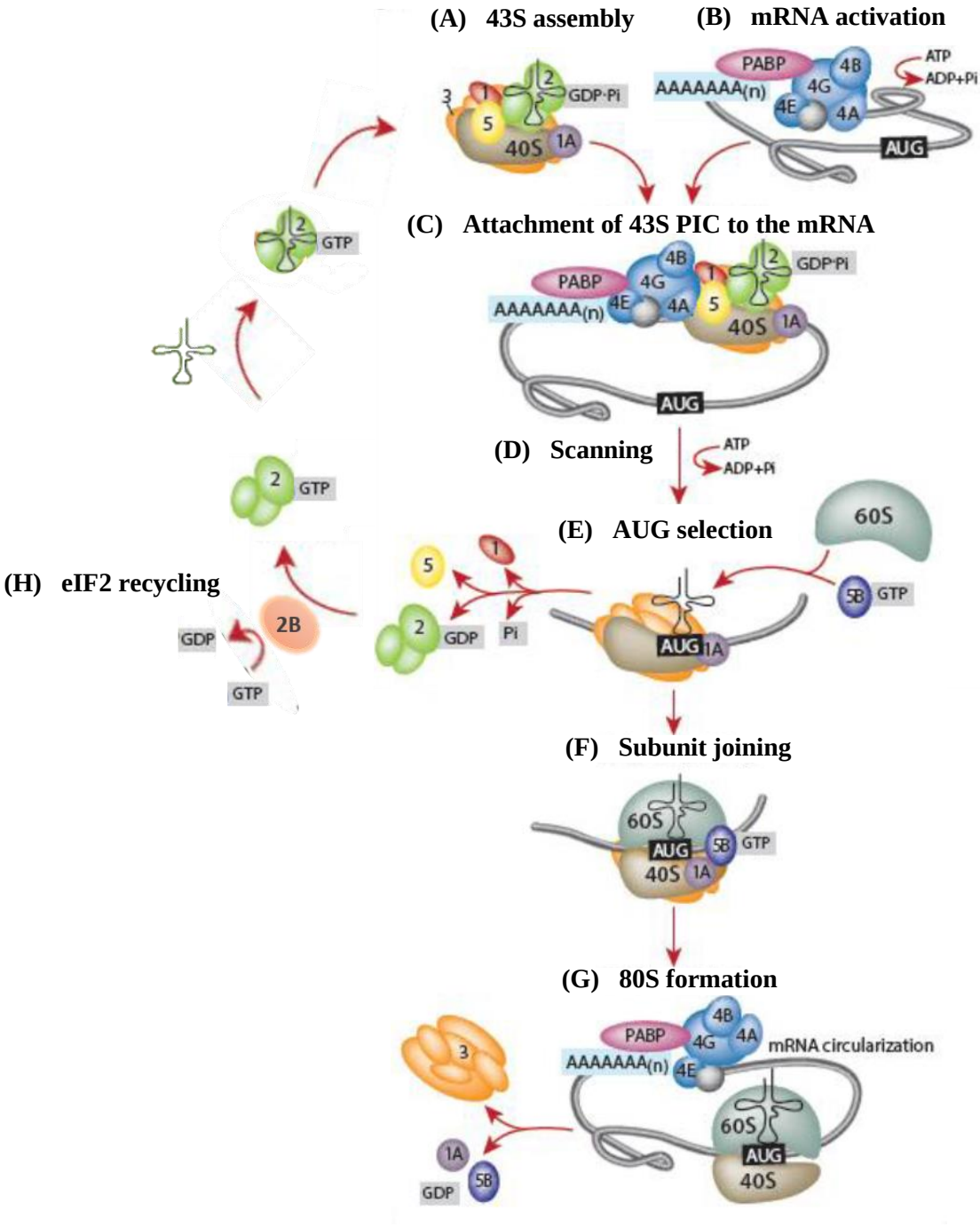


Figure 1. 2 — Schematic representation of the canonical cap-dependent translation initiation. (A) 43S assembly: After formation of the ternary complex, composed of eukaryotic initiation factor (eIF) 2 bound to both initiator methionyl transfer RNA (Met–tRNA$_i$) and guanosine triphosphate (GTP), assembly of the 43S preinitiation complex (PIC) occurs. Here, the ternary complex binds to the 40S ribosomal subunit with the aid of a multifactor complex composed of eIF1, eIF1A, eIF3 and eIF5. **(B) mRNA activation:** Before attachment to the 43S complex, the mRNA must be activated. This is accomplished by the eIF4F complex (eIF4E, eIF4G and eIF4A) and the poly(A)-binding protein (PABP). eIF4F recognizes, through eIF4E subunit, the cap structure at the 5'end of the mRNA. Then, the cap-proximal region is unwound in an adenosine triphosphate (ATP)-dependent manner by eIF4A subunit, which is aided by eIF4B and eIF4G. **(C) Attachment of 43S PIC to the mRNA:** After mRNA activation, PIC and mRNA attachment is mediated by eIF4G and eIF3 interactions. **(D) Scanning:** After being assembled onto the mRNA, the 48S complex scans the 5' UTR base by base until an AUG in a good context is selected through codon–anticodon base pairing with the Met–tRNA$_i$. **(E) AUG selection:** Complementation with the anticodon of the Met–tRNA$_i$ results in scanning arrest and release of eIF1, eIF2, phosphate (Pi) and eIF5 from the PIC. **(F) Subunit joing:** The assembly of the 60S ribosomal subunit is accomplished by the GTPase eIF5B and the eIF1A factor. **(G) 80S formation:** After subunit joining, eIF5B hydrolyses GTP and leaves the initiation complex, as well as the eIF1A initiation factor. The 80S complex is, then, formed. **(H) eIF2 recycling:** The guanosine nucleotide exchange factor eIF2B recycles eIF2, through guanosine diphosphate (GDP) replacement with GTP, ensuring another translation cycle. (Adapted from Haimov *et al.*, 2015.[16])

1.1.1.2 Translation elongation

The translation elongation cycle begins with the anticodon of the Met–$tRNA_i$ at the peptidyl (P) site of the ribosome, base-pared with the start codon.[23] The second codon of the open reading frame (ORF) is at the acceptor (A) site of the ribosome, waiting for the correspondent aminoacyl-tRNA, which is delivered as part of a ternary complex composed of GTP and the eukaryotic elongation factor (eEF) 1A.[20,23,24] Although ternary complexes with noncognate aminoacyl–tRNAs can bind to the A-site of the ribosome, steps involving codon–anticodon base pairing between the mRNA and tRNA, and conformational changes in the decoding center of the ribosome guarantee that only the cognate tRNA is selected for the next stage of elongation.[20] Codon recognition by the tRNA triggers GTP hydrolysis by eEF1A, allowing the release of this factor and the accommodation of the aminoacyl-tRNA into the A-site.[23] After that, the peptidyl bond formation between the incoming amino acid and the peptidyl-tRNA occurs.[23] This reaction is catalyzed by the peptidyl transferase center of the large ribosomal subunit.[20,23,24] The peptidyl bond formation triggers the movement of the ribosomal subunits, which place the deacylated tRNA in a hybrid state, with its acceptor end at the exit (E) site of the ribosome and its anticodon end at the P-site.[20,23] The peptidyl tRNA is also in a hybrid situation.[20,23] Its acceptor end is at the P-site and its anticodon end is at the A-site.[20,23] Translocation of the deacylated tRNA and the peptidyl tRNA to the canonical E and P sites, respectively, requires GTP hydrolysis by the GTPase eEF2.[20] This translocation also allows the placement of the next mRNA codon at the A-site and another cycle of elongation begins (Figure 1.3)[23]. All the steps described above are repeated until a stop codon appears, triggering the termination process.[20,23]

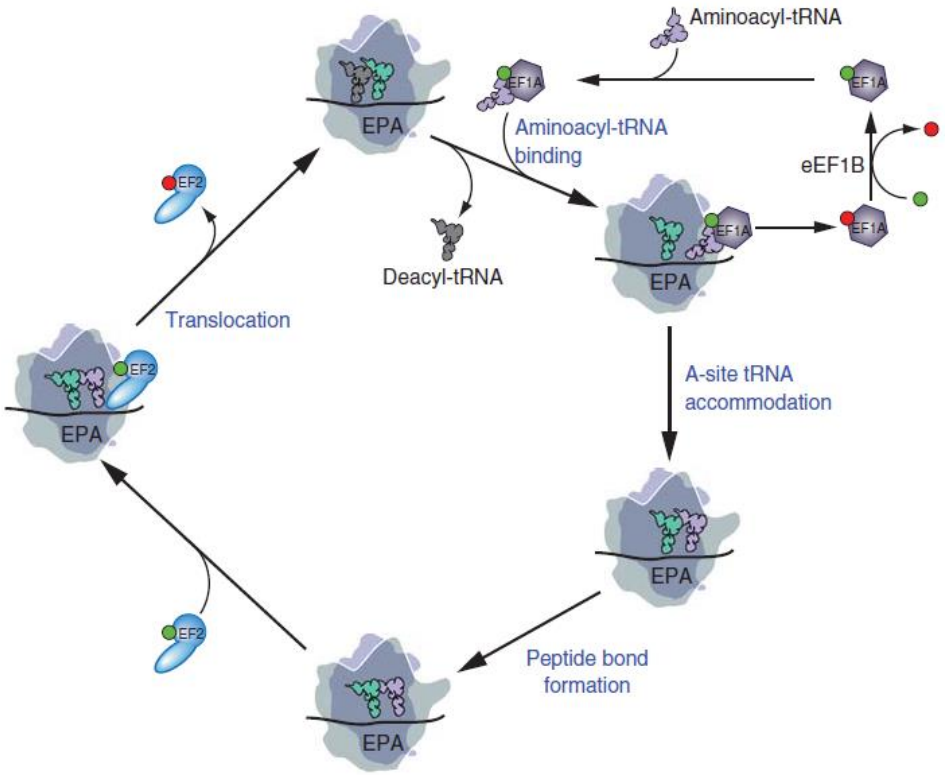


Figure 1. 3 — Schematic representation of the translation elongation cycle. The eukaryotic elongation factor (eEF) 1A–guanosine triphosphate (GTP)–aminoacyl–transfer RNA (tRNA) ternary complex delivers the cognate aminoacyl–tRNA to the acceptor (A) site of the small subunit. After codon–anticodon base pairing, eIF1A–GTP is hydrolyzed to eEF1A–GDP (GDP, guanosine diphosphate) and the aminoacyl–tRNA is accommodated into the A-site. eEF1A–GDP is recycled to eEF1A–GTP by the guanine nucleotide exchange factor eEF1B. At the ribosome, peptidyl bond formation occurs. It triggers the movement of the ribosomal subunits, which place the deacylated tRNA in a hybrid state, with its acceptor end at the exit (E) site of the ribosome and its anticodon end at the peptidyl (P) site. The peptidyl tRNA is also in a hybrid situation. Its acceptor

end is at the P-site and its anticodon end is at the A-site. Binding of eEF2–GTP promotes translocation of the tRNAs into the canonical P and E sites. This event is followed by release of eEF2–GDP, which unlike eEF1A does not require an exchange factor. This translocation also allows the placement of the next mRNA codon at the A-site and another cycle of elongation begins. (Adapted from Dever and Green, 2012.[23])

Each cycle of the translation elongation is followed by a recycling event. After GTP hydrolysis by eEF1A and subsequent dissociation from the ribosome, the eEF1A–GDP must be recycled to its GTP-bound form.[20,23] This is accomplished by the eEF1B, which catalyzes the guanine nucleotide exchange on eEF1A.[20,23]

1.1.1.3 Translation termination and recycling

Translation termination occurs when the last codon of the ORF—UAA, UGA or UAG stop codon— enters the A-site of the ribosome.[20,23,25] In this process, the ester bond between the polypeptide chain and the P-site tRNA must be hydrolyzed by the peptidyl transferase center of the large ribosomal subunit to allow the release of the newly synthetized polypeptide.[20] Two protein factors, the eukaryotic release factors (eRFs) 1 and 3, assist eukaryotic translation termination (Figure 1.4).[23] eRF1 is a tRNA-shaped protein factor composed of three domains that recognizes the stop codon and stimulates nascent peptide chain release.[20,23,25] Its N-terminus domain has conserved sequences that are responsible for codon recognition by codon–anticodon-like interactions.[23,25] On the other hand, the middle domain plays a crucial role in polypeptide–tRNA ester bond hydrolysis, because it contains a universally conserved glycine–glycine–glutamine (Gly–Gly–Gln) motif that interacts with the peptidyl transferase center.[23,25] Finally, the C-terminus mediates the association between eRF1 and eRF3.[23,25] The latter is a translational GTPase that accelerates peptide release and increases translation termination efficiency in a GTP-dependent manner.[23] Mechanistically, upon recognition of a stop codon, the eRF1–eRF3–GTP ternary complex binds to the A-site of the ribosome and GTP hydrolysis occurs.[23] This leads to accommodation of the middle domain of eRF1 in the peptidyl transferase center, which is followed by polypeptide–tRNA bond hydrolysis and peptide chain release.[23] After GTP hydrolysis, eRF3 leaves the ribosome.[23] At this stage, the 80S ribosome is still bound to the now deacylated tRNA, to the mRNA, and possibly to the eRF1.[20,23] Thus, the next step involves recycling of the translational machinery, which is characterized by the dissociation of the ribosomal subunits and subsequent release of the mRNA and the deacylated tRNA.[20,23] However, the mechanisms underlying this process are still largely unkown.[20,23]

Recent studies revealed that the adenosine triphosphate (ATP)-binding cassette subfamily E member 1 (ABCE1) is a likely candidate for promoting ribosomal recycling.[23] Some models propose that ABCE1 converts the chemical energy from ATP hydrolysis into mechanical motions that can separate the ribosomal subunits.[23] Besides the role in recycling, ABCE1 also seems to contribute to polypeptide release during translation termination.[23] The deacylated tRNA and mRNA are dissociated from the isolated small subunit during a process enhanced by Ligatin (also known as eIF2D).[23]

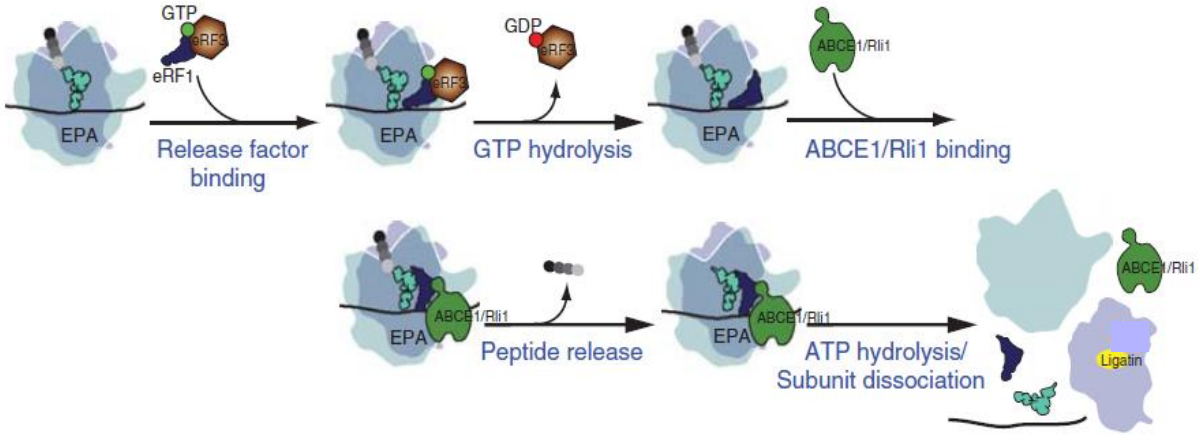


Figure 1. 4 — Schematic representation of translation termination and recycling. Translation termination, characterized by stop codon recognition and subsequent polypeptide release, is assisted by the eukaryotic release factors eRF1 and eRF3. eRF1 is a tRNA-shaped protein factor composed of three domains that recognizes the stop codon and stimulates nascent peptide chain release. Regarding eRF3, it is a translational GTPase that accelerates peptide release and increases translation termination efficiency in a guanosine triphosphate (GTP)-dependent manner. Mechanistically, upon recognition of a stop codon, the eRF1–eRF3–GTP ternary complex binds to the A-site of the ribosome and GTP hydrolysis occurs. This is followed by polypeptide–tRNA bond hydrolysis and peptide chain release. The recycling of the translational machinery is done by the adenosine triphosphate (ATP)-binding cassette subfamily E member 1 (ABCE1). This protein also seems to contribute to polypeptide release during translation termination. The deacylated tRNA and mRNA are dissociated from the isolated small subunit during a process enhanced by Ligatin. (Adapted from Dever and Green, 2012.[23])

1.1.2 Non-canonical translation initiation mechanisms

Although many eukaryotic mRNAs are only translated through the canonical scanning pathway, some mRNAs can also initiate translation via alternative mechanisms.[16,26] Non-canonical translation initiation mechanisms described so far can be cap-dependent and 5' UTR scanning-independent, cap-independent and 5' UTR scanning-dependent or cap and 5' UTR scanning-independent.[16,26,27] For instance, mRNAs with an extremely short 5' UTR are dependent on the 5'-end cap to initiate translation, nevertheless they avoid the scanning mechanism.[16] In this case, translation initiation is mediated by the TISU (translation initiator of short 5' UTR) element, which contains the start codon in the middle of its sequence and is located close to the cap structure.[16] Another cap-dependent but scanning-independent translation initiation mechanism is the ribosome shunting, in which the ribosome bypasses/shunts parts of the 5' UTR on its way to the AUG start codon.[16] mRNAs with highly stable secondary structures in the 5' UTR use this mechanism to bypass these obstacles and initiate translation.[16] One example of a non-canonical translation initiation mechanism that is cap-independent is the cap-independent translational enhancer (CITE)-mediated translation.[27] Although this is a cap-independent mechanism, it remains dependent on 5'-end scanning.[27] It involves special elements—CITEs—that are located within both mRNA untranslated regions.[27] They interact with key initiation factors in order to promote the assembly of translation initiation complexes.[27] The internal ribosome entry site (IRES)-mediated translation is also a cap-independent mechanism of translation initiation.[26,27] It will be described in detail in the next paragraphs.

1.1.2.1 Internal ribosome entry site-mediated translation

Studies on viral gene expression were essential for the discovery of this mechanism.[27–29] The *poliovirus* and *encephalomyocarditis virus* mRNAs were the first to be described to use IRES

elements.[28] This alternative mode of translation initiation is generally independent of the 5'-end cap recognition and allows the direct recruitment of the 40S ribosome to the vicinity of the initiation codon.[27]

Analysis of the structures of viral IRES elements has shown they possess complex secondary and tertiary structures that interact with components of the canonical translational apparatus, allowing the assembly of the translational machinery through a 5'-end independent manner.[29] Moreover, there are some cases in which translation initiation proceeds without the need of canonical factors, relying entirely on interactions between the IRES and the 40S ribosome.[29] The IRES-mediated translation initiation can also be assisted by other protein factors that are not involved in the cap-dependent translation (Figure 1.5).[27,29] They are called IRES *trans*-acting factors (ITAFs), and will be described later in detail.[27,29]

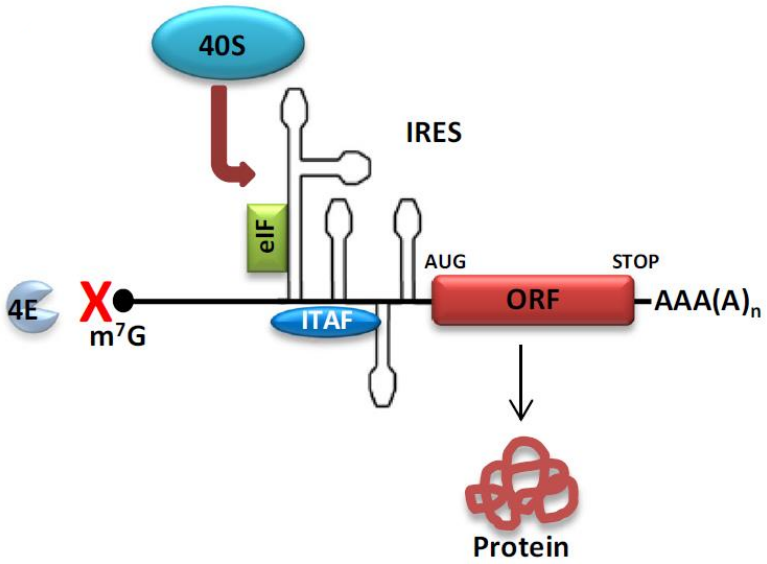


Figure 1. 5 — Schematic representation of internal ribosome entry site (IRES)-mediated translation initiation. This alternative mode of translation initiation is mediated by strong mRNA secondary structures and is generally independent of 5'-end cap recognition. It allows the direct recruitment of the 40S ribosome to the vicinity of the initiation codon (AUG) of the open reading frame (ORF). IRES-mediated translation initiation can be assisted by some canonical initiation factors (eIFs) and/or by other auxiliary proteins, named IRES *trans*-acting factors (ITAFs). (Adapted from Lacerda et al., 2017.[27])

Many viral IRES elements share primary sequence or have secondary structure similarity.[30] For instance, *picornavirus* (*PCV*) and *hepatitis C virus* (*HCV*) IRESs show common sequence motifs at similar positions, such as the UGGG sequence around loop regions.[31] Furthermore, regarding viral IRESs secondary and tertiary structures and their need for protein factors, four groups can be presented.[32,33] In group I, IRESs can bind directly to the ribosome, without the aid of canonical factors and/or ITAFs.[33] The Met–$tRNA_i$ is also not required.[33] This group is composed by highly structured IRESs that fold into a very compact structure.[32] The class II comprises the IRES elements that use some canonical factors, such as eIF3 and eIF2, and need the Met–$tRNA_i$ to initiate translation.[33] In this case, the IRES element maintain some packed and structured regions.[32] Extended and largely flexible IRESs constitute classes III and IV of viral IRES elements.[33] In both classes, besides some canonical factors, other auxiliary proteins (ITAFs) are needed.[33] While in class III translation is initiated at the IRES, in class IV the initiation of translation occurs at an AUG codon downstream of the IRES.[33]

Several eukaryotic cellular mRNAs can also be translated in an IRES-dependent manner.[29] Like viral IRES-containing mRNAs, cellular mRNAs containing IRES elements have a reduced need for canonical initiation factors and/or specific ITAFs.[29] However, unlike viral IRESs, cellular IRESs are less structured, with no common motif being described.[29] As far as cellular IRESs primary sequence is

concerned, some common features, such as U and pyrimidine (C/U) rich regions, where recently identified as strong determinants of IRES activity.[34]

In 2016, performing systematic mutagenesis on IRES elements, Weingarten-Gabbay *et al*. divided both viral and cellular IRESs into two functional classes.[35] One functional class comprises the IRESs in which expression is reduced only when a specific position is mutated.[35] The other contains the IRESs in which mutations in most positions significantly reduce protein expression.[35] For the first functional class, IRESs can act through a short sequence motif, for example, an ITAF-binding site, in which only mutations in this specific motif reduce activity.[35] For instance, the IRES of the runt related transcription factor 1 (RUNX1), a factor that regulates the differentiation of hematopoietic stem cells into mature blood cells, is included in this class.[35,36] In the second class, IRES can form a secondary structure, in which mutations at various positions can disrupt the overall structure, reducing IRES activity.[35] It is the case of the *cationic amino acid transporter 1* (*cat-1*) and the *avian erythroblastosis virus* (*AEV*) IRESs.[35] However, this study presents some limitations. For instance, they only used a short length IRES library, so they could not include long IRESs on their systematic approach.[35] Therefore, their results and conclusions are limited to short IRESs.[35,37] Besides, other experiments validating this information have not yet been performed, pointing out the lack of knowledge that still exists concerning IRES elements.[26,37]

Focusing now on the cellular IRESs, it is well known that protein synthesis consumes a lot of the cell energy and resources.[27,38] Therefore, under unfavorable or energy-depriving conditions, such as endoplasmic reticulum (ER) stress, hypoxia, nutrient starvation, mitosis and cell differentiation, the cell must unsure only vital processes in order to canalize the energy for stress recovery or for the energy-depriving situation.[27,38] To do that, canonical translation is impaired and translation of specific mRNAs only occurs through non-canonical translation mechanisms, such as IRES-mediated translation.[27,29] Many may be the mechanisms by which cap-dependent translation is inhibited.[30] For instance, one of them, involving the mammalian target of rapamycin (mTOR) signaling pathway, can prevent the cap-binding protein eIF4E from forming the eIF4F complex, thus inhibiting translation initiation.[39,40] mTOR is a serine/threonine kinase that plays a key role in controlling cell growth and metabolism, by sensing nutrient availability, cellular stress and the energy status of the cell.[39] This enzyme can regulate protein translation through eIF4E-binding proteins (4E-BPs).[40] These proteins have a conserved amino acid motif, also present in eIF4G, that can bind to eIF4E.[39,40] Thus, 4E-BPs can compete with eIF4G for eIF4E binding.[40] 4E-BPs can also alternate between a hypophosphorylated and a phosphorylated state.[39,40] Under unfavorable growth conditions such as stress and nutrient deprivation, with mTOR low activity, 4E-BPs are not phosphorylated.[39] In this state, 4E-BP compete with eIF4G for eIF4E binding, which prevents the formation of the eIF4F complex.[39] When mTOR activity is high, at favorable growth conditions, mTOR phosphorylates 4E-BPs, leading to 4E-BP dissociation from eIF4E and consequent protein translation.[39] Since eIF4E acts as a rate-limiting factor of cap-dependent translation because of its low cell availability, this mechanism is an efficient way to impair canonical translation.[39] Other central mechanism of canonical translation initiation inhibition during stress is the eIF2α subunit phosphorylation.[30] Phosphorylated eIF2α binds to eIF2B, which prevents its GDP–GTP exchange activity needed for eIF2 recycling.[30] Without the conversion of eIF2–GDP to eIF2–GTP, the formation of the ternary complex and consequently canonical translation are inhibited.[30] There are four eIF2α kinases: the general control nonderepressible 2 (GCN2) kinase, the protein kinase R (PKR), the PKR-like endoplasmic reticulum kinase (PERK) and the heme-regulated eIF2α kinase (HRI).[41] Each of them is active under specific stress conditions.[41] While GCN2 phosphorylates eIF2α under nutrient deprivation, eIF2α phosphorylation is accomplished by PERK during ER stress.[41] Regarding HRI and PKR, the former is active in erythroid cells during heme deprivation and the latter phosphorylates eIF2α during viral infection.[41] Repression of canonical translation is then followed by translation of mRNAs that encode crucial proteins for cell survival and stress recovery (for instance, apoptotic proteins, heat-

shock proteins, tumor suppressors, cyclins, transcription factors, receptors, channels, growth factors and transporters).[27,29] Not all IRESs are active when cap-dependent translation is inhibited.[27] In fact, depending on the stress and what caused the impairment of cap-dependent translation, only selected IRES-containing mRNAs will be translated.[27,29] Some examples will be presented later, in the paragraph about the physiological importance of IRES-mediated translation.

Besides IRES-mediated translation under stress conditions, cellular IRESs have also been implicated in various physiological processes such as spermatogenesis, neuron plasticity and cell differentiation.[27] Indeed, under normal physiological conditions with cap-dependent translation fully active, they allow the expression of IRES-containing mRNAs that have highly structured 5' UTRs, which is incompatible with efficient scanning.[29] Besides that, there are also mRNAs that can be translated by the cap-dependent mechanism and by an IRES element.[42] It is the case of neurogranin, a neuronal calmodulin-binding protein that is expressed in dendrites.[42] These cells have relatively low levels of eIF4E, and therefore cap-dependent translation is not efficient in ensuring the expression of all transcribed mRNAs, including neurogranin.[42] Therefore, IRES elements can also translate these mRNAs that are not efficiently translated by the canonical translation.[42]

Thus, cellular IRESs play two major roles in the cell.[27,29] In normal physiological conditions, they assist translation of mRNAs that do not accomplish the necessary protein levels only by cap-dependent translation.[42] On the other hand, under cap-dependent translation inhibitory conditions, IRES-mediated translation ensures the production of a selective group of proteins necessary to respond to the stress or condition that triggered canonical translation initiation impairment.[27] Besides that, while in IRES-mediated translation under stress, the protein production is robust, in normal physiological conditions, IRES-mediated translation usually produces low protein levels.[29]

1.1.2.1.1 IRES *trans*-acting factors

As mentioned before, there are non-canonical factors, the ITAFs, that can participate in IRES-mediated translation.[27,29,30] Although being widely studied, the mechanisms underlying ITAF function are still not known.[27,29] However, some theories have been proposed: (1) ITAFs can remodel IRES spatial structures to produce conformations with higher or lower affinity for the components of the translational machinery; (2) they can be used to complement the action of the canonical factors, thus functioning as adaptors between the mRNA and the ribosome, allowing or not the interaction between these two counterparts; (3) they can take the place of the canonical factors as mediators of mRNA and ribosome interactions.[27,29] Besides that, several studies suggest that the subcellular (nucleus/cytoplasm) distribution of ITAFs is also fundamental to determinate IRES activity.[29] However, the importance of ITAF's compartmentalization in IRES-mediated translation is not fully understood.[29] Once again, different hypotheses exist.[29] One of them suggests that ITAFs associate with IRES elements in the nucleus, sequestering the mRNAs in that cell compartment, thus inhibiting translation.[29] The other postulates that ITAFs are in their unbound form in the nucleus while the target IRES-containing mRNAs are in the cytoplasm.[29] Following the appropriate signals, either the ITAF-bound mRNAs (in the first theory) or the unbound ITAFs (in the second theory) translocate from the nucleus to the cytoplasm, allowing the translation of the mRNAs via an IRES-dependent manner.[29]

One of the most studied ITAFs is the polypyrimidine-tract-binding (PTB) protein, an RNA-binding protein that is involved in translation and in mRNA splicing, stability and localization within the cell.[43] This ITAF seems to be required for the function of many cellular IRESs under apoptotic conditions, thus allowing the expression of apoptotic mRNAs when cap-dependent translation is impaired in part due to 4E-BP and eIF4E binding.[44] For instance, together with another ITAF, the

upstream of N-Ras (Unr)[ii], PTB regulates the expression of apoptotic protease activating factor 1 (Apaf-1).[45] First, when Unr recognizes its binding site on *Apaf-1* IRES, located within a structural domain, this ITAF unwinds that region.[45] These changes of structure in *Apaf-1* IRES allow the interaction of *Apaf-1* IRES with PTB, which in turn also disrupt RNA–RNA interactions, facilitating ribosome recruitment.[45] The same occurs with the IRES of the anti-apoptotic protein Bcl-2-associated athanogene 1 (Bag-1), which also suffers structure remodeling due to RNA chaperone activities of PTB and poly(rC)-binding protein (PCBP).[45] Moreover, PCBP has also been associated with IRES-mediated translation of *c-myc*, a transcription factor that plays an important role in cellular proliferation, formation and differentiation.[46,47]

Other known described ITAFs belong to the group of heterogeneous nuclear ribonucleoproteins (hnRNPs), known to shuttle between the nucleus and the cytoplasm.[27,29,30] They play an important role in pre-mRNA processing as well as in mRNA export, stability and translation.[30] Some examples of ITAFs from this group include hnRNP A1, hnRNP C1/C2, hnRNP I, hnRNP E1/E2, hnRNP K and hnRNP L.[27] For example, hnRNP C1/C2, together with the La autoantigen and other auxiliary proteins, can form a protein complex that functionally interacts and regulates *X-linked inhibitor of apoptosis protein* (*XIAP*) IRES. [27,29,30]

However, ITAFs do not always work as activators of IRES-mediated translation.[48] For instance, PTB works as a repressor of binding immunoglobulin protein (BiP) internal initiation, through BiP 5' UTR binding.[48] An ITAF can then be a positive or negative regulator of IRES-mediated translation.[45]

Canonical factors can also be involved in IRES-mediated translation.[27,29] However, the need for canonical factors seems to vary among different cellular IRESs containing mRNAs, even between those from the same functional family.[29] For instance, while c-*myc* and N-*myc*[iii] IRESs need eIF4A and eIF3 for their activity, without requiring eIF4E or eIF4G, L-*myc*[ii] IRES necessitates both eIF4E and full-length eIF4G for its action.[49] Moreover, although these three IRESs require the ternary complex (eIF2–GTP–Met–$tRNA_i$) for their function, N-*myc* IRES seems to be less affected by the reduction of the amount of eIF2–GTP–Met–$tRNA_i$ in the cell.[49] Besides that, the role of eIF2 has also been investigated for IRES-mediated translation of other cellular IRES-containing mRNAs.[29] Many of them (*cat-1*, for example) were shown to be insensitive to the inhibition of protein synthesis caused by eIF2α phosphorylation.[29] Consequently, these IRES-containing mRNAs must use different mechanisms for Met–$tRNA_i$ delivery.[29] It is thought that, in these cases, eIF5B canonical factor and/or Ligatin play an important role in promoting Met–$tRNA_i$ and ribosome binding.[29] However, a study performed in 2011 by Thakor and Holcik showed that some cellular IRES (*cellular inhibitor of apoptosis protein* (*cIAP*), *Apaf-1* and *death-associated protein 5* (*DAP5*) IRES) do not need eIF5B to initiate translation.[50] Thus, other mechanisms of Met–$tRNA_i$ delivery may exist.[51] Nevertheless, this was a limited investigation and other studies are required to confirm and explain other Met–$tRNA_i$ delivery mechanisms.[51]

All in all, although the information about IRES-mediated translation initiation is scarce, the known data indicates that different cellular IRESs have very different requirements for canonical initiation factors and ITAFs.[29] Besides, even though canonical factors are involved in some cases of IRES-mediated translation, this translation initiation mechanism is not canonical, since the cap structure is not needed to recruit the ribosome.[52] In addition, different cellular IRESs reveal different responses to various stress conditions.[29] For instance, during mitosis, the *Unr* and the c-*myc* IRES become more active, while others do not.[29] Additionally, during apoptosis, the *Apaf-1* IRES is active while the *XIAP* IRES is inhibited.[29] Therefore, further research in this area is required to a better understanding of IRES-mediated translation regulation.[27]

[ii] Cytoplasmic RNA-binding protein, involved in mRNA stability.[150]
[iii] Transcription factor closely related to *c-myc*.[47]

1.2 Physiological and pathophysiological significance of IRES-mediated translation

Most of the cellular IRESs are found in mRNAs whose protein products are needed for cell growth and cell death control.[45,53] Hence, it has become clear that internal initiation is an important cellular mechanism implicated in the regulation of gene expression and with high significance in physiological/stress conditions.[29,45] In fact, it is well known that cells can be exposed to certain physiological conditions during which cap-dependent protein synthesis is reduced, but mRNAs that encode for proteins involved in cell adaptation to cellular stress continue to be translated.[45,53]

One example that illustrates the importance of IRES-mediated translation is the case of limited amino acid availability in mammalian cells. Amino acid starvation induces phosphorylation of eIF2α, leading to a global decrease in protein synthesis.[29,45] This induces the transcription of amino acid transporter genes, such as *CAT-1*, and its subsequent IRES-mediated translation.[29,45] Thus, the cells are prepared to transport amino acids once they are available.[29] IRES-mediated translation has also been shown to play an important role in the regulation of gene expression during mitosis.[29] For instance, the cyclin-dependent kinase 11 (CDK11), which is involved in spindle formation, is translated in an IRES-dependent way at the G2/M transition, when cap-dependent translation is impaired.[29,45] Efficient translation of several other proteins, such as c-*myc* and apoptotic inducer factors, is also maintained by IRES-mediated translation during apoptosis.[45] Besides, in hypoxia, cells can also preserve the translation of, for example, vascular endothelial growth factor (VEGF) and hypoxia inducible factor 1 alpha subunit (HIF-1α) in an IRES-dependent manner.[45] Briefly, after being translated, HIF-1α can promote the expression of several genes, including VEGF coding gene, in order to protect and adapt the cell to low oxygen levels.[54] In turn, VEGF expression will promote angiogenesis and nitric oxide synthase activity.[55] This enzyme produces nitric oxide, responsible for vasodilatation and increased blood flow.[55] These adaptor responses intent to augment the oxygen levels in the cell, promoting cell survival.[55]

Nevertheless, IRES-mediated translation can also benefit pathological conditions. Indeed, the presence of IRES elements in so many genes involved in regulation of cell fate and survival raises the possibility that deregulated internal initiation may contribute to tumorigenesis. This dark side of the IRES-mediated translation will be described below.

1.2.1 IRES-mediated translation and cancer

In a tumor's microenvironment, characterized by cellular unfavorable conditions, such as nutrient deprivation, hypoxia, genotoxic and oxidative stress, cap-dependent translation is usually impaired.[27,53] As stated before, these inhibitory events of canonical translation are followed by cap-independent translation of specific proteins needed for cell survival and stress recovery, which in malignant cells is often synonymous of IRES-mediated translation of key regulatory proteins that promote tumor proliferation and cancer progression.[27,53]

Growth factors are extremely important in these events.[27,56] Fibroblast growth factors (FGF) are essential for proliferation and differentiation of a wide variety of cells and tissues.[27] Some of them, such as FGF1 and FGF2, contain IRES elements within their 5' UTRs that allow their expression when canonical translation is impaired.[27] FGF2 translation is considered a critical step in tumorigenesis, since its expression has been implicated not only in solid tumors, but also in multiple myeloma.[27] Moreover, the platelet-derived growth factor (PDGF), which is critical in regulation of mesenchymal cell migration and proliferation, has a B form that is regulated by an IRES element, whose aberrant expression has already been implicated in cancer progression.[56]

Alternatively, due to the quick proliferating rate of malignant cells, growth factor deprivation can also occur.[27] This induces IRES-mediated translation of specific transcripts, such as the XIAP mRNA.[27]

XIAP is the most potent intrinsic inhibitor of caspases, thus it can protect tumor cells from apoptosis.[27,56] In addition, IRES-mediated translation of XIAP can difficult cancer treatments.[27] In fact, non-canonical translation of this anti-apoptotic factor is increased in response to radiation, which makes tumor cells resistant to radiotherapy.[27,56] Other anti-apoptotic proteins are translated via IRES elements during oxidative and genotoxic stress caused by chemotherapy.[27] It is the case of Bag-1, which can promote tumor cells resistance to etoposide and arsenite treatments.[27] Thus, IRES-mediated translation of pro-survival factors can also provide malignant cells with mechanisms for chemotherapy and radiotherapy resistance.[27]

ITAFs can also play crucial roles in cancer progression.[57] Their subcellular localization is usually deregulated in tumor cells, which leads to aberrant translation through IRES mechanisms, thus promoting tumor growth, cell cycle progression, resistance to apoptosis, angiogenesis and cell migration.[57]

1.3 IRES as therapeutic targets

As mentioned before, IRES-mediated translation of specific mRNAs may contribute to the development of severe diseases and pathological states, such as cancer.[27,29] Because of that, these mRNA elements began to be studied with a therapeutic objective.[58] Approaches in this field are focused on the design of antagonists/drugs that can disrupt the IRES itself or prevent IRES interactions with the ribosome and with auxiliary factors necessary for their function.[58] These approaches include, for instance, antisense oligonucleotides and small-molecule inhibitors.[58] The antisense oligonucleotides can act through two distinct mechanisms: either guiding the destruction of the IRES through an RNAse-dependent manner or preventing IRES interactions with the ribosome.[58] The possibility of inhibiting *HCV* IRES using antisense oligonucleotides by both approaches has already been demonstrated.[59] Moreover, focusing now on the small-molecule inhibitors, a few compounds have recently shown the ability to also suppress the *HCV* IRES.[58] These compounds also inhibit some cellular IRES known to be involved in carcinogenesis, such as *c-myc* and *VEGF*.[58] The exact mechanism of action of small-molecule inhibitors is not well understood, but it is believed that these drugs act as IRES intercalating agents, thus preventing the binding between the mRNA and the ribosome.[58]

Nevertheless, all these therapeutics have some challenges, such as the difficulty in achieving efficient delivery of the therapeutic molecules to the target cells, which highlights the need for further investigation in this field.[58]

From another perspective, IRES elements can be used for the design of multicistronic vectors, as in bacterial operons, in order to promote the translation of several genes using the same mRNA (Figure 1.6).[60] A concrete example is the bicistronic IRES-based vector, expressing FGF2 and cysteine-rich angiogenic inducer 61 (Cyr61), used in a mice hindlimb ischemic muscle model to stimulate angiogenesis.[61] The use of this vector proved to be more efficient than the use of the monocistronic vectors expressing each factor alone.[61] Besides, this bicistronic vector did not show side effects in promoting angiogenesis, unlike the monocistronic vectors, which is very important to ensure safety in clinical assays.[61]

Most IRES-based vectors developed up to now use viral IRESs, because they show stronger efficiency in transient transfections, when compared to cellular IRESs.[60] Indeed, cellular IRESs show low efficiency in transiently transfected cells, in part due to cell and tissue specificity of these cellular RNA elements.[60] Nevertheless, this concept of translational tissue-specificity may be applied in therapeutics by coupling tissue-specific IRESs with tissue-specific promoters to create vectors with increased safety.[60] This means that with cellular IRESs, an adequate IRES could be chosen according to the cell type or tissue to be targeted, thus reducing secondary effects and augmenting specificity.[60] Since IRES-mediated translation is enhanced in stress conditions, the IRES-based multicistronic vectors could

be used to treat pathological conditions characterized by cellular unfavorable environments, such as cancer and ischemic diseases.[60]

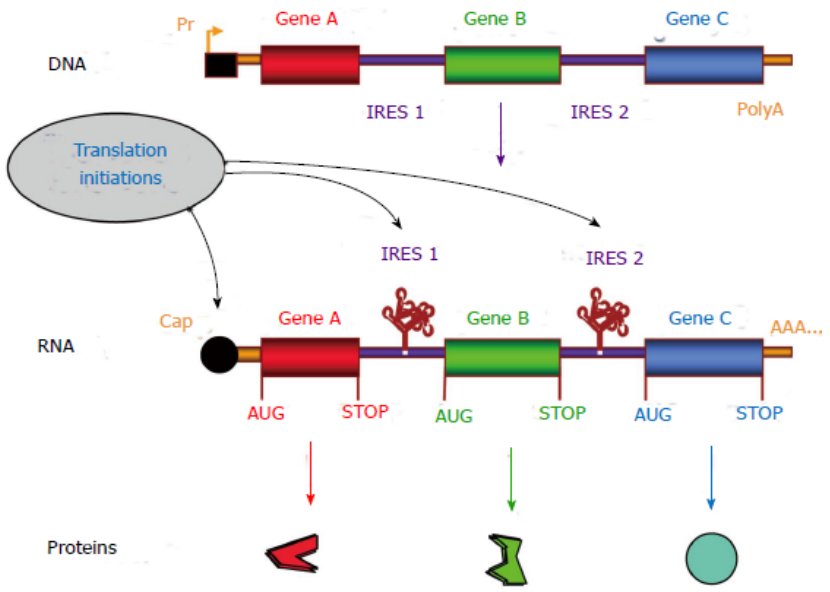


Figure 1. 6 — Internal ribosome entry site (IRES)-based multicistronic vector. IRES-based vector contains several genes, separated by IRESs, that are under the control of the same promoter (Pr). With this transcription unit, a single mRNA can codify different genes. Translation initiation occurs at the 5'-end by the cap-dependent mechanism, resulting in translation of the first open reading frame (Gene A). Internal translation initiations occur at each IRES, resulting in translation of genes B and C (Adapted from Renaud-Gabardos *et al*., 2015.[60])

1.4 IRES-mediated translation and p53 – their combined role in cancer

1.4.1 p53 protein

p53 is a tumor suppressor protein that plays a key role in maintaining genome integrity and preventing malignant transformation.[62–64] As a transcription factor, p53 regulates different biological processes, such as cell cycle, apoptosis, stem cell differentiation, senescence and DNA repair, not only by activating protein-coding genes, but also for allowing the expression of many miRNAs.[65] This protein can also regulate different aspects of metabolism, including glycolysis, mitochondrial oxidative phosphorylation, fatty acid synthesis and oxidation.[63] Usually, p53 is inactive in the cell and its levels are low.[62,63] Low protein levels are kept by p53 negative regulators, such as Hdm2, the murine double minute 2 (Mdm2) human homolog, which is an E3 ubiquitin ligase that degrades p53 through the proteasome degradation pathway.[63] In turn, p53 can regulate Hdm2 expression, which means that high levels of p53 can induce Hdm2 production.[66] So, these two proteins act within a negative feedback loop.[66] A significant increase in Hdm2 levels, after the induction of p53 expression, promotes p53 degradation and restores p53 homeostatic levels.[66] The reduction of p53 levels is then followed by a decrease in Hdm2 levels.[66] p53 expression can be triggered by DNA damage, nutrient deprivation, hypoxia and oncogene activation.[63,64] These signals lead to p53 activation, a process that comprises three steps: p53 stabilization, sequence-specific DNA binding and transcription of the target genes.[67] In p53 stabilization, p53 ubiquitination by Hdm2 is inhibited, thus preventing its degradation.[67] For sequence-specific DNA binding, p53 must be acetylated at specific amino-acid residues.[67] Then, interaction of p53 with transcription factors and other p53 posttranslational modifications (for instance, phosphorylation and methylation) allow the expression of the target genes.[67] Depending on cell and

tissue types, the type and intensity of the stress signal, p53 can induce cell cycle arrest, apoptosis or senescence.[63,64] While transient stresses with reparable damages lead to a p53 survival response, prolonged stresses with irreversible damages trigger p53-dependent cell death pathways.[65] Nevertheless, all the outcomes intent to promote tumor suppression.[65]

1.4.1.1 p53 isoforms

p53 can distinctly regulate many cellular pathways in part because it has many isoforms, expressed under specific cellular conditions, that have different roles in the cell.[68] In fact, *TP53* gene encodes for more than twelve p53 isoforms that can be generated by alternative splicing, internal initiation of translation and transcription from an internal promotor located in intron 4 of *TP53* gene (Figure 1.7A).[69] The canonical p53 protein (FL-p53 or p53α), composed of seven functional domains, is the most abundant p53 isoform.[69] Its N-terminus encompasses two transactivation domains, TAD1 and TAD2.[69] A proline-rich domain (PXXP) and a DNA-binding domain (DBD) constitute the middle of p53 protein.[69] The C-terminus has a nuclear localization signaling domain (NLS), an oligomerization domain (OD) and a negative-regulation domain (Neg).[69] All other p53 isoforms have the lack of one or more functional domains (Figure 1.7B).[69]

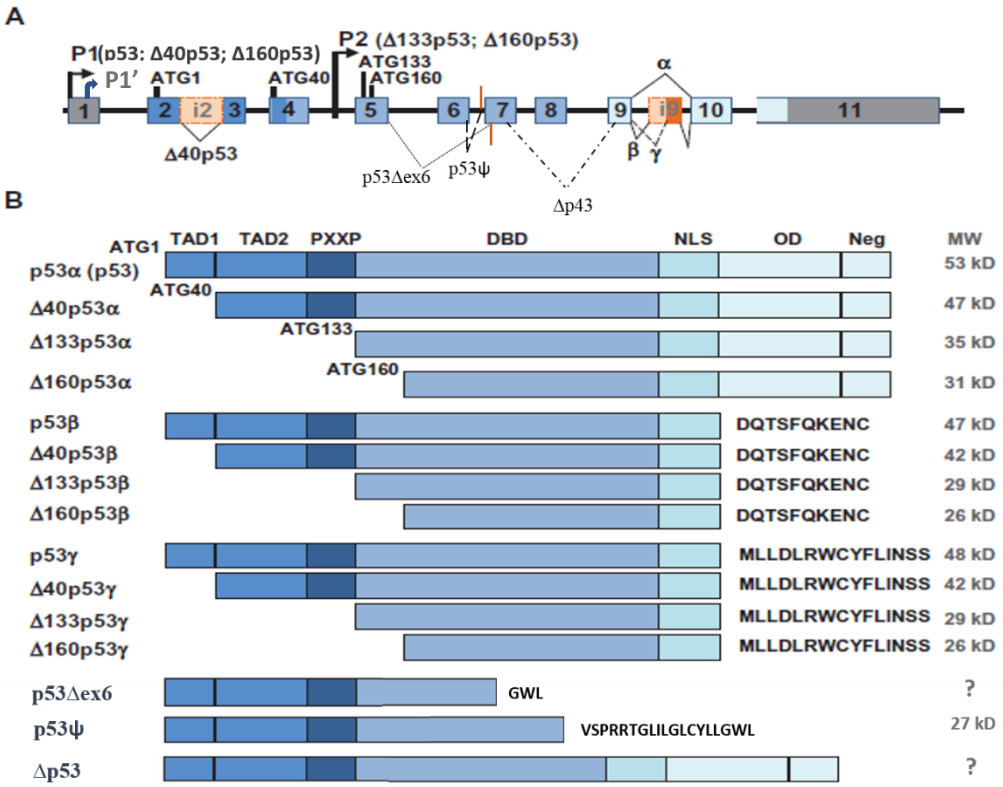


Figure 1. 7 — TP53 gene (A) and p53 protein isoforms (B). For *TP53* gene, composed of eleven exons (represented by numbers from 1 to 11), fifteen p53 protein isoforms have already been described. p53α is the most abundant p53 protein isoform and has all seven functional domains: TA1 and TA2 – transactivation domain 1 and 2, respectively; PXXP – proline-rich domain; DBD – DNA-binding domain; NLS – nuclear localization signaling domain; OD – oligomerization domain; and Neg – negative-regulation domain. The other p53 isoforms can be generated by internal promotor existence, alternative splicing and internal translation initiation. P1 and P1' correspond to proximal promoters and P2 corresponds to an internal promotor. The premature STOP codons that appear after alternative splicing are represented by ▮. i2 and i9 represents introns 2 and 9, which are involved in generation of some p53 isoforms by alternative splicing. MW – molecular weight, kD – kilodalton (Adapted from Surget *et al*., 2013.[69] Construction of p53ψ, Δp43 and p53Δex6 was performed considering information from Senturk *et al*., 2014[70], Rohaly *et al*., 2005[71] and Pekova *et al*., 2008.[72])

TP53 gene contains two promoters upstream exon 2 (P1 and P1' at Figure 1.7A) and an internal promoter at intron 4 (P2 in Figure 1.7B).[73,74] Both promoters P1 and P1' produce full-length *p53* transcripts that allow the expression of FL-p53, Δ40p53 and Δ160p53 protein isoforms.[69,73,75]The internal promoter produces a shorter transcript that promotes the expression of Δ133p53 and also Δ160p53.[68,69,75] Besides, splicing in intron 9 can produce three different C-terminal domains (α, β and γ).[69] Altogether, twelve p53 isoforms can be generated: FL-p53 or p53α, p53β, p53γ, Δ40p53 or Δ40p53α, Δ40p53β, Δ40p53γ, Δ133p53 or Δ133p53α, Δ133p53β, Δ133p53γ, Δ160p53 or Δ160p53α, Δ160p53β and Δ160p53γ.[69] Three less studied p53 isoforms have also been reported: p53ψ, created by alternative splicing in intron 6 with a premature STOP codon generation[70]; Δp53, obtained by junction of exon 7 with exon 9[71]; and p53Δex6, a truncated p53 protein, due to the appearing of a premature STOP codon, after alternative splicing with ORF frameshift.[72]

1.4.2 p53 and IRES-mediated translation

As mentioned before, translation of some p53 isoforms can occur through internal initiation, more specifically, by IRES-mediated translation.[68,69] For instance, FL-p53 and Δ40p53 isoforms expression can be accomplished by IRES elements (Figure 1.8A).[68] While FL-p53 can also be translated by the cap-dependent mechanism, Δ40p53 translation is only ensured by this alternative mechanism of translation initiation.[53,62,64] IRES-mediated translation of both isoforms is enhanced under different stress conditions, including DNA damage, ionizing radiation, ER stress and cancer.[27,64] It has been demonstrated that both *FL-p53* and Δ*40p53* IRESs are active during the G2/M transition, when cap-dependent translation is impaired.[64,68] Interestingly, at G1/S transition, only the Δ*40p53* IRES is functional.[64,68] Moreover, Thapsigargin-induced ER stress triggers unfold protein response and specifically increases Δ40p53 isoform over the FL-p53.[64] Therefore, expression of different p53 protein isoforms can be differential, depending on the stimulus.[64]

Δ*40p53* IRES is located between the two initiation codons existing at the full-length transcripts.[68] This IRES element contains a hairpin structure that seems to interact with Hdm2, PTB and hnRNP C.[68] These three proteins are activators of Δ40p53 translation and have been shown to work during stress conditions.[68] Therefore, they can be considered Δ*40p53* IRES ITAFs.[68] In fact, it is known that, following DNA damage, PTB translocates from the nucleus to the cytoplasm, and increased abundance of PTB contributes to Δ*40p53* IRES activity.[62] This protein can also induce the *FL-p53* IRES, but with less extent.[62] *FL-p53* IRES activity can also be modulated, for instance, by PTB-associated splicing factor (PSF), DAP5 and translational control protein 80 (TCP80).[27,62] Curiously, the p53 protein itself seems to function as a negative ITAF on its own IRES.[68,76]

More recently, an IRES element regulating Δ160p53 isoform expression was also identified, but little information exists about it (Figure 1.8A and 1.8B).[77] It is known that this IRES allows the translation of Δ160p53 isoform and is located downstream AUG160 start codon.[77] Indeed, Δ160p53 coding region up to nucleotide 432 was proposed as the putative Δ*160p53* IRES, whose activity is stimulated under stress conditions.[77] Furthermore, it has been reported that Δ*160p53* 5' UTR has an inhibitory effect on this IRES element.[77] Nevertheless, further investigation on Δ*160p53* IRES is required to unveil the mechanisms underlying non-canonical expression of this shorter p53 isoform. Which proteins regulate its expression? Are there canonical factors involved in IRES-mediated translation of Δ160p53 or translation is accomplished only by non-canonical auxiliary proteins? How is the secondary structure of Δ*160p53* IRES? These are some of the questions that need to be addressed.

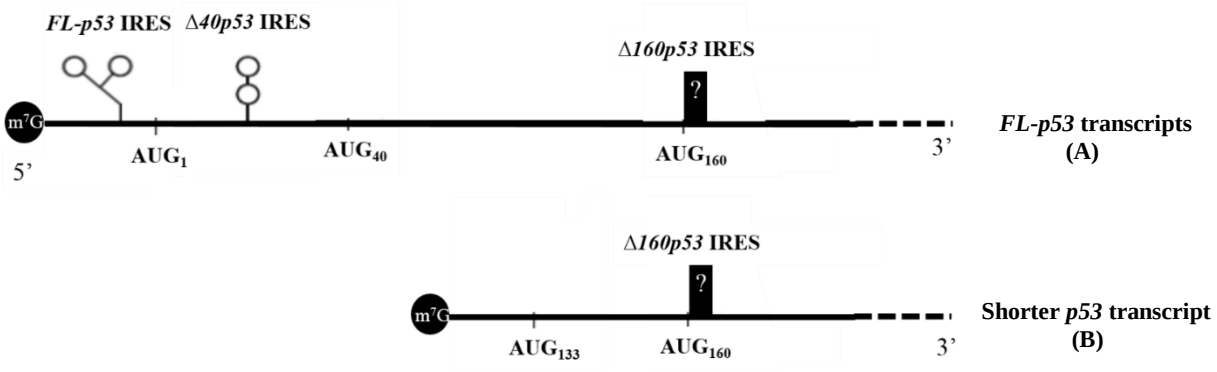


Figure 1. 8 — Internal ribosome entry sites (IRESs) of *p53* transcripts. (A) The full-length *p53* transcripts allow the expression of FL-p53, Δ40p53 and Δ160p53 protein isoforms. FL-p53 can be translated by the cap-dependent mechanism or through *FL-p53* IRES. On the other hand, Δ40p53 and Δ160p53 translation is only accomplished through Δ*40p53* and Δ*160p53* IRESs, respectively. **(B)** The shorter *p53* transcript allows the expression of Δ133p53 and Δ160p53 isoforms. Δ133p53 expression is only accomplished through the canonical cap-dependent translation initiation mechanism. Regarding Δ160p53, its expression is mediated by an IRES element located downstream AUG_{160} start codon. ^{m7G} represents the 7-methylguanosine cap structure and ? indicates Δ*160p53* IRES, whose secondary structure is still unknown. (Adapted from Candeias, 2011.[68])

1.4.3 p53 isoforms and cancer

p53 protein is frequently mutated in cancer.[63] Unlike many other tumor suppressors, which are frequently inactivated by deletions or truncating mutations in tumors, p53 mutations are usually missense mutations that lead to the production of a full-length mutant protein.[63] Therefore, in addition to the loss of tumor suppressor activity, tumor-associated mutant p53 proteins often gain new tumorigenic activities, named gain-of-functions (GOF).[63,75] Many of these oncogenic functions of mutant p53 include promotion of cell proliferation, angiogenesis, migration, invasion, metastasis, chemoresistance and metabolic changes that promote tumor cell survival.[63] For instance, studies in mutant p53 knock-in mouse models showed that mice expressing p53 with R172H mutation, the equivalent R175H hotspot mutation in humans, developed an altered spectrum of tumors.[78,79] Moreover, it has also been reported that mice expressing R270H mutant p53 (R273H mutant p53 in human) developed tumors with a high metastatic capacity.[79]

Mutations on *TP53* gene are usually within p53 DNA-binding domain.[80] Although some p53 mutants can bind to the DNA and, thus, directly control gene expression, it is though that the main mechanism of action of these mutant p53 proteins is through the binding to other transcription factors.[80] For instance, the transactivation domain-containing p63 isoform (TAp63), a transcription factor from p53 family, can bind to mutant p53 but not to wild-type p53.[80] This interaction leads to TAp63 inactivation, which triggers the transcription of genes associated with pro-invasive features.[80] Nevertheless, either by direct regulation of gene expression or by interaction with other transcription factors, research on this field suggests that mutant p53 can influence most of the hallmarks of cancer (for instance, evasion from tumor suppressors and from the immune system, resistance to apoptosis and angiogenesis induction).[81]

From other perspective, an imbalance among different p53 isoform expression is also observed.[69] For instance, while in normal breast tissues, FL-p53, p53β and p53γ are expressed, in 60 % of breast tumors, expression of p53β and p53γ is abrogated. Moreover, overexpression of Δ133p53 is usually reported.[69] In fact, recent studies suggest that Δ133p53 may have pro-oncogenic traits.[82] Mouse models expressing Δ122p53 (murine homolog to human Δ133p53) developed a wide range of tumors, such as lymphoma and osteosarcoma, characterized by cellular reduced apoptosis and increased proliferation.[82]

Besides, in colon tumors, transition from adenoma to carcinoma is followed by an increase in Δ133p53 levels.[82]

As for p53β, despite its impaired expression in more than half of breast cancers, this p53 isoform can also promote tumorigenesis.[69,82] For instance, its expression is increased in renal cell carcinoma and colon adenoma.[69,82]

Δ40p43 also has an ambiguous role in cancer.[64] On the one hand, it can act alone or in combination with FL-p53 to promote apoptosis and consequent tumor suppression.[64] On the other hand, Δ40p53 is highly expressed in melanoma cells, whereas in normal melanocytes its expression is almost null.[82]

For Δ160p53, until recently, little was known about its role on carcinogenesis.[27,75] Now, some evidence attribute pro-oncogenic functions to this shorter p53 isoform.[27,75] Indeed, Δ160p53-overexpressing cells showed to exhibit cancer phenotypes, such as enhanced cell survival, proliferation, invasion and adhesion.[27,75] Moreover, R273H mutant p53 has been shown to induce Δ160p53 overexpression, which highlights that p53 mutations can induce the imbalance of p53 isoform expression.[75] Furthermore, as stated before, this p53 isoform is expressed through an IRES element whose activity is stimulated under stress.[77] So, as tumor microenvironment is characterized by stress conditions, non-canonical translation of this shorter p53 isoform may have a preponderant role in cancer.[77] However, to confirm these assumptions, more studies are required.

2. Aims of this study

Under unfavorable and energy-depriving conditions that impair cap-dependent translation, IRES-mediated translation assists the synthesis of key proteins involved in stress response, cell growth and cell death control.[27,29,38] As already stated, many mRNAs containing IRES elements are deregulated in cancer, thus contributing to cancer progression.[27,53] As this mechanism of translation initiation seems to play a key role in promoting tumorigenesis, the understanding of how IRES-containing mRNAs are regulated under different stress conditions and how the switch between cell homeostasis and cell neoplastic transformation is triggered is of great matter.

p53 is a tumor suppressor protein that plays a key role in maintaining genome integrity and preventing malignant transformation.[65] As a transcription factor, p53 can regulate different biological processes in part because it has many isoforms, expressed under specific cellular conditions, that have different roles in the cell.[65] Recently, Δ160p53 isoform overexpression was implicated in survival, proliferation and invasion of tumor cells.[75] Since expression of this shorter p53 isoform is regulated by an IRES element[77], we hypothesized that Δ*160p53* IRES is crucial for Δ160p53 tumorigenic functions.

Therefore, we aimed to investigate the mechanisms that regulate IRES-mediated translation of tumor suppressor p53 isoform Δ160p53, under stress conditions known to impair cap-dependent translation or mutation statuses known to induce oncogenic functions in p53. Hence, we proposed to address:

1) The effect of stress conditions that are known to inhibit cap-dependent translation in Δ*160p53* IRES function and Δ160p53 protein isoform expression;

2) The impact of p53 cancer mutations in Δ*160p53* IRES activity and Δ160p53 protein isoform expression;

3) The identification of novel IRES *trans*-acting factors (ITAFs) that regulate Δ160p53 protein isoform expression.

In parallel, we also aimed to identify new cancer-related mRNAs non-canonically translated during stress conditions, with the aim of gaining an insight on IRES-mediated reprogramming of cellular function during stress and carcinogenesis.

3. Materials and Methods

3.1 Plasmid constructs

The bicistronic constructs were based on psiCHECK™–2 vector (Promega). This vector has two reporter genes: *Renilla* Luciferase (RLuc) and *Firefly* Luciferase (FLuc). While RLuc is cap-dependent translated, FLuc is only translated through cap-independent mechanisms. The empty vector, *PR_F*, in which a stable hairpin was cloned downstream RLuc stop codon to prevent reinitiation[77], the negative control for cap-independent translation, *PR_HBB_F*, with the 5' UTR of human β-globin cloned after the stable hairpin and upstream FLuc sequence[83], and the positive control for cap-independent translation, *PR_MYC_F*, with *c-myc* IRES sequence[83], were already available in the laboratory. The same occurred with the promoterless empty vector (*P-R_F*)[83] and the bicistronic constructs used as positive controls for cryptic promoter activity assessment: *PR_MLH1_F*[83], the one containing *MLH1* 5' UTR, and its promoterless counterpart, *P-R_MLH1_F*[83].

The following Δ160p53 bicistronic constructs were also already available to use: *PR_d160_F*[77] (containing Δ*160p53* IRES—Δ160p53 coding sequence up to nucleotide 432); *PR_5'd160_F*[77] (containing Δ*160p53* IRES and its 5' UTR—Δ*160p53* IRES + Δ133p53 coding sequence up to Δ*160p53* start codon, excluding Δ133p53 AUG start codon); *P-R_d160_F*[iv] (promoterless biscistronic construct with Δ*160p53* IRES); *P-R_5'd160_F*[iii] (promoterless biscistronic construct with Δ*160p53* IRES and its 5' UTR); *PR_5'd160R175H_F*[84], *PR_5'd160R248Q_F*[84] and *PR_5'd160R273H_F*[84] (bicistronic constructs containing Δ*160p53* IRES and its 5' UTR with the hotspot mutations R175H, R248Q and R273H, respectively).

In order to obtain *PR_5'd160R282W_F* bicistronic construct, containing Δ*160p53* IRES and its 5' UTR with R282W missense mutation, site-directed mutagenesis with NZYProof DNA Polymerase (NZYTech) was performed, according to standard procedures. To generate R282W mutation, primers #1 and #2 were used (Table 3.1). The bicistronic construct *PR_5'd160_F* was the template. After site-directed mutagenesis PCR, the construct was digested with *DpnI* (Roche) for 1 h at 37 °C and purified using DNA Clean & Concentrator -5™ Kit (Zymo Research), according to manufacturer's instructions. Afterwards, *E. coli* NZY5α competent cells (NZYTech) were transformed with the above-mentioned constructs. Bacteria were cultivated in Luria-Bertani (LB) agar plates supplemented with 100 µg/mL of ampicillin, overnight at 37 °C with shaking [220 rotations per minute (rpm)]. Plasmid DNA was obtained through overnight liquid cultures of single colonies, followed by DNA extraction with NZYMiniprep kit (NZYTech), according to manufacturer's instructions. Mutation confirmation, using primer #3, was accomplished by Sanger sequencing method (Table 3.1). RLuc and FLuc sequence integrity was also confirmed by Sanger sequencing method [primers #4 to #15 (Table 3.1)].

To obtain the monocistronic constructs *pFLp53_12MS2*, *p5'd133p53_12MS2*, *p5'd160p53_12MS2* and *p5'd160IRES_12MS2*, composed of, respectively, full-length p53, Δ133p53 coding sequence with part of its 5' UTR[v], Δ160p53 coding sequence with its 5' UTR[vi] and Δ160p53 IRES[vii] with its 5' UTR[viii], fused with twelve *MS2* repeats, four strategies of cloning were performed. The empty vector with the twelve *MS2* repeats (*p_12MS2*) was already available in the laboratory.[ix] In the first cloning strategy, p53 sequences were amplified using primers containing linkers for *EcoRI*

[iv] Construct obtained in Doctor Luísa Romão's lab, Instituto Nacional de Saúde Dr. Ricardo Jorge, Lisbon, Portugal.
[v] Last 21 nucleotides of the 5' UTR.
[vi] Δ133p53 coding sequence up to AUG160, excluding Δ133p53 start codon.
[vii] Δ160p53 coding sequence up to nucleotide 432.
[viii] Δ133p53 coding sequence up to AUG160, excluding Δ133p53 start codon.
[ix] Construct obtained in Doctor Carmo Fonseca' lab, Instituto de Medicina Molecular, Lisbon, Portugal.

(forward primers) and *BglII* (reverse primers). PCR reaction was done with NzyProof DNA polymerase (NZYTech), according manufacturer's instructions, with primers #16 and #17, #18 and #17, #19 and #17, #19 and #20 (Table 3.1) for full-length p53, Δ133p53 coding sequence with part of its 5' UTR, Δ160p53 coding sequence with its 5' UTR and Δ160p53 IRES with its 5' UTR amplification, respectively. A pcDNA3.1 vector containing full-length p53 coding sequence was the template. After p53 sequence amplifications, PCR products were purified using DNA Clean & Concentrator -5™ Kit (Zymo Research), following manufacturer's instructions. Afterwards, both PCR products and *p_12MS2* empty vector were digested, first, with *EcoRI* (Amersham) for 1 h at 37 °C, and then, with *BglII* (NZYTech), overnight, at 37 ˚C. For the second cloning strategy, only *BglII* was used. Primers #21 and #17, #22 and #17, #23 and #17, #23 and #20 (Table 3.1), containing linkers for this enzyme were used for full-length p53, Δ133p53 coding sequence with part of its 5' UTR, Δ160p53 coding sequence with its 5' UTR and Δ160p53 IRES with its 5' UTR amplification, respectively. The template was the same as before. Purified PCR products and *p_12MS2* vector were, then, digested with *BglII* for 1 h at 37 °C. *p_12MS2* vector was also subjected to an alkaline phosphatase (NZYTech) treatment, to avoid plasmid re-ligation, for 1 h, at 37 °C. In both strategies, digestion products were separated in a 1 % agarose gel and purified using NZYGel Pure extraction kit (NZYTech), according to manufacturer's instructions. Afterwards, digested fragments were ligated with T4 DNA Ligase (NZYTech) for 3 h at 23 ˚C, using a vector/insert molar ratio of 1:3. Thereafter, *E. coli* NZY5α competent cells (NZYTech) were transformed with the above-mentioned constructs. Bacteria were cultivated in LB agar plates supplemented with 100 µg/mL of ampicillin, overnight at 37 °C with shaking (220 rpm). Then, a colony screening PCR was performed, using GoTaq® Flexi Polymerase (Promega) and following manufacturer's instructions. The primers used for p53 sequence of interest amplification were used in this screening. Randomly selected colonies were transferred to a 10 µl water-containing tube and 1 µl from this mixture was added to the colony screening PCR solution. Positive colonies detected by the colony screening PCR were cultured overnight in liquid medium supplemented with 100 µg/mL of ampicillin at 37 °C while shaking (220 rpm). Plasmid DNA was extracted using NZYMiniprep kit (NZYTech), following manufacturer's instructions. To assess cloning success, Sanger sequencing method was performed, using primer #24 (Table 3.1). The third and the fourth cloning strategies are very similar. In both cases, p53 sequence amplifications were accomplished following the same protocol as in strategy two. Then, while PCR products were digested for 5 h at 37 °C with *BglII*, *p_12MS2* empty vector was, first, digested with *BsrGI* (New England BioLabs) overnight at 37 °C and, then, with *BglII* for 5 h, at the same temperature. Digested fragments were purified by gel extraction, as before. Vector digestion with *BsrGI* and *BglII* produced two fragments: one with 4,540 bp and another one with 649 bp. In the third cloning strategy, p53 sequences and 4,540 bp fragment ligation, with a 1:3 vector/insert molar ratio was performed. Then, a second ligation was performed to connect the aforementioned ligation product with the 649 bp fragment. A vector/insert molar ration of 1:3 was also used. In the fourth cloning strategy, ligation in a molar ratio of 1:1 was performed to link p53 sequences with the 649-bp fragment. Then, ligation between the obtained fragment and the 4,540-bp fragment was performed. In all cases, T4 DNA ligase was used, as described before. The rest of the cloning protocol is also the same as the one used for the two first strategies. In some cases, after colony screening PCR and after plasmid DNA extraction from positive colonies, a screening by digestion of positive plasmids with *BsrGI* and *BamHI* was done, for 1 h at 37 °C, before sequencing.

The monocistronic construct *pΔN6_HDM2*[x] (with the human Mdm2 homolog HDM2), was already available in the laboratory.

All digestions and ligations were done following manufacturer's instructions.

[x] Construct obtained in Doctor Robin Fåhraeus' former lab, University of Dundee, Dundee, United Kingdom.

All DNA sequencing was performed using Big Dye™ Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems™), according to manufacturer's instructions.

Plasmid DNA quantification was accomplished using NanoDrop™ (Thermo Scientific™) apparatus, by measuring the absorbance at 260 nanometers (nm).

Table 3.1 – Sequences of the primers used to produce or to confirm the integrity of the described constructs

Primer	*Sequence (5' → 3')*
#1	TGTGCCTGTCCTGGGAGAGACTGGCGCACAGAGGAAGAGAATC
#2	GATTCTCTTCCTCTGTGCGCCAGTCTCTCCCAGGACAGGCACA
#3	GGACGCTCCAGATGAAATGG
#4	GTCTCGAACTTAAGCTGCAG
#5	ATGGCTTCCAAGGTGTACGA
#6	GAGAACGCCGTGATTTTTCTG
#7	CAGGAGAGGGTAGGCCGTCTA
#8	ATCCCTCTCGTTAAGGGAGGC
#9	CGAAGTACTCGGCATAGGTG
#10	CACCTATGCCGAGTACTTCG
#11	GTGAGAGAAGCGCACACAG
#12	GCAAATCAGGTAGCCCAGG
#13	CTGTGTGCGCTTCTCTCAC
#14	CCGCCTGAAGTCTCTGATCA
#15	CCGCCCCGACTCTAGAATTA
#16	CCGGAATTCATGGAGGAGCCGCAGTC
#17	GGAAGATCTTCAGTCTGAGTCAGGCCCTTCT
#18	CCGGAATTCTACTCCCCTGCCCTCAAC
#19	CCGGAATTCTTTTGCCAACTGGCCAAGACC
#20	GGAAGATCTGCTCCCTGGGGGCAGCTC
#21	GGAAGATCTATGGAGGAGCCGCAGTC
#22	GGAAGATCTTACTCCCCTGCCCTCAAC
#23	GGAAGATCTTTTTGCCAACTGGCCAAGACC
#24	GTTGACGCAAATGGGCGGTAG

3.2 Cell culture and plasmid transfection

HeLa, A549 and HEK-293T cells were cultured in Dulbecco's modified Eagle's medium (1 x DMEM + GlutaMAX™-I; Gibco® by Life Technologies™) supplemented with 10 % volume/volume (v/v) fetal bovine serum (FBS; Gibco® by Life Technologies™). They were incubated at 37 °C in an atmosphere containing 5 % CO_2.

Cells were split in 24 and 6-well plates, in 60-mm and 100-mm dishes. Cells in 24-well plates were transfected 24 h post-seeding, at 70–90 % cell confluence, with 375 ng of plasmid DNA, using 293fectin Transfection Reagent (Invitrogen® by Life Technologies™) and Reduce Serum Medium (Opti-MEM® I; Gibco® by Life Technologies™), following manufacturer's instructions. For 6-well plates, 60-mm and 100-mm dishes, volumes were four, ten and twelve times greater, respectively. When co-transfections were performed, 6-well plates and 750 ng of each plasmid DNA were used.

3.3 Drug treatment and cell lysis

For luminometry assays, HeLa cells were treated, 24 h post-transfection, with 1, 2 or 4 µM Thapsigargin (Sigma-Aldrich) for 16 h. In parallel, cells were treated with the corresponding control vehicle, DMSO [0.1 % (v/v)]. 16 h after drug treatment, cells were rinsed with 1 x cold phosphate-buffered saline (PBS) and lysed with 1 x passive lysis buffer (Promega) (20 µL for 100 % cell confluence). Cell lysates were rapidly frozen at -80 °C, until they were used for luminometry assays. For cryptic promoter activity and alternative splicing assessment, transfected but untreated cells were rinsed with 1 x cold PBS 24 h post-transfection. While for cryptic promoter activity evaluation, cells were lysed with 1 x passive lysis buffer (100 µL for 100 % cell confluence), followed by lysate storage at -80 °C until luminometry assays, for alternative splicing control, cells were lysed with RA1 buffer [Nucleospin® RNA II extraction kit (Macherey-Nagel)], followed by RNA extraction.

For Hdm2-RNA co-immunoprecipitation (co-IP) experiments, two different immunoprecipitation (IP) buffers were used: IP buffer A [1 % (v/v) Nonidet P-40 (Invitrogen), 150 mM NaCl, 20 mM Tris-HCl pH 7.4, supplemented with 1:100 Protease Inhibitor Cocktail (Sigma-Aldrich)]; and IP buffer B [1 % (v/v) Nonidet P-40 (Invitrogen), 100 mM NaCl, 10 mM $MgCl_2$, 10 % (v/v) glycerol, 50 mM Tris-HCl pH 7.5, supplemented with 1:100 Protease Inhibitor Cocktail (Sigma-Aldrich)]. For co-IP experiments with IP buffer A, 48 h post-seeding (without transfection) or 24 h post-transfection, 100 % confluent cells were rinsed with 1 x cold PBS and then harvest with 1 mL 1 x cold PSB into a microcentrifuge tube. Cells were centrifuged at 6,000 g for 6 min in a tabletop refrigerated microcentrifuge (Eppendorf 5415R), the medium was removed and the cell pellet stored at -80 °C. Cell pellets were thawed at 37 °C and lysis was performed using 300 µL of IP buffer per 100-mm dish, for 40 min, on ice, while vortexing every 5 min. When IP buffer B was used, 48 h post-seeding (without transfection) or 24 h post-transfection, cells were rinsed with 1 x cold DMEM + GlutaMAX™-I (Gibco® by Life Technologies™). Then, they were lysed and collected with 450 µL of IP buffer B, followed by centrifugation for 10 min at 2,300 g, in a tabletop refrigerated microcentrifuge (Eppendorf 5415R), and pellet removal.

When MG132 was used, cells were treated with either 0.1 % (v/v) DMSO (vehicle) or 25 µM of MG132 (Calbiochem), 4 h prior to lysis with IP buffer B.

For the experiment in which lysis efficiency of IP buffers was tested, volumes were three times less. After lysis with IP buffer A, lysates were centrifuged 15 min at maximum speed in a tabletop refrigerated microcentrifuge (Eppendorf 5415R), and 20 µL of supernatant was transferred to a tube containing 5 µL 5 x (v/v) SDS sample buffer (NZYTech) and stored at -20 °C. For lysis with IP buffer B, 20 µL of the final supernatant was also transferred to a tube containing 5 µL 5 x (v/v) SDS sample buffer and stored at -20 °C. Cells lysed with 1 x SDS sample buffer were first rinsed with 1 x cold PBS followed by addition of 1 x SDS sample buffer (100 µL of 1 x SDS sample buffer for 100 % cell confluence). Lysates were collected to a microcentrifuge tube and stored at -20 °C.

3.4 Luminometry assays

Cell lysates were thawed at 37 °C, followed by centrifugation at maximum speed for 4 min in a tabletop microcentrifuge (Eppendorf, Centrifuge 5415R). Cleared lysates (10 µL) were then used, following manufacturer's instructions, to perform the luminometry assays in a GloMax® 96 microplate Luminometer (Promega). The Dual Glo Assay System (Promega) was used to determinate RLuc and FLuc relative luciferase activity as well as β-galactosidase activity. Briefly, to assess FLuc and RLuc luciferase activities, 40 µL of Luciferase Assay Reagent (LAR), containing the substrate for firefly luciferase, was added to the sample and luminescence was read in the GloMax® 96 microplate Luminometer (Promega). Then, 40 µL of Stop & Glo Reagent, which stops reaction between LAR and

firefly luciferase and contains the substrate for *Renilla* luciferase, was added to the sample and luminescence was read. Regarding β-galactosidase activity determination, 50 µL of Beta Glo Reagent was added to 10 µL of sample, followed by an incubation of 45 min at room temperature. Then, bioluminescence signals were read in the GloMax® Luminometer. Beta Glo Reagent provides a coupled enzyme reaction system utilizing a luciferin-galactoside substrate. This substrate is cleaved by β-galactosidase to form luciferin and galactose. The luciferin is then utilized in a firefly luciferase reaction to generate light.

All results were obtained in arbitrary light units.

3.5 Hdm2 immunoprecipitation

When IP buffer A was used, on the day before total lysate preparation, 35 µL of rProtein G Agarose Beads (Invitrogen™) per sample were washed three times with 1 x cold PBS, followed by centrifugation for 1 min at 5,900 g, at 4 °C. Then, G Beads were resuspended with 280 µL of IP buffer A and incubated overnight at 4 °C, in a spinning rotator, with either 1:100 dilution SMP14 or 4B2 mouse anti-Hdm2 monoclonal antibodies (gift from Doctor Roman Hrstka, Masaryk Memorial Cancer Institute, Brno, Czech Republic). Incubation of G beads with no antibody was also done as a control (MOCK). On the following day, cell lysates were centrifuged in a tabletop refrigerated microcentrifuge (Eppendorf 5415R) at maximum speed for 15 min, at 4 °C. The supernatant was collected to a new microcentrifuge tube. Then, 30 µL of total supernatant was transferred to a microcentrifuge tube containing 7.5 µL of 5 x SDS sample buffer and stored at -20 °C (total lysate sample for Western blot analysis). The remaining lysate was evenly split, depending on the number of samples. Afterwards, total lysates were incubated with the G-beads, overnight at 4 °C, in a spinning rotator. On the next day, G-beads were centrifuged at 5,900 g for 1 min, at 4 °C, and 30 µL of supernatant was collected to a new microcentrifuge tube with 7.5 µL of 5x SDS sample buffer and stored at -20 °C (flow-through sample for Western blot analysis). The remaining supernatant was discarded and beads were washed three times with IP buffer A, sequentially diluted to half with 1 x cold PBS, and each time centrifuged at 5,900 g for 1 min, at 4 °C. After the last wash and complete supernatant removal, 20 µL of 2 x SDS sample buffer was added to the beads. The IP sample was also stored at -20 °C, until Western blot analysis.

For Hdm2 immunoprecipitation with IP buffer B, 60 µL of total lysate was transferred to a tube with 15 µL of 5 x SDS sample buffer and stored at -20 °C (total lysate sample for Western blot analysis). Then, 500 µL of supernatant, per sample, was incubated overnight at 4 °C, in a spinning rotator, with either 1:100 dilution SMP14 or 4B2 mouse anti-Hdm2 monoclonal antibodies. Incubation of G beads with no antibody was also done as a control (MOCK). On the following day, 100 µL of G beads (Invitrogen™) per sample were washed three times with IP buffer B, followed by centrifugation at 2,300 g for 30 s, in a tabletop refrigerated microcentrifuge (Eppendorf 5415R). After the last wash, 200 µL of IP buffer B were added to the tube in order to resuspend the beads, which were then incubated with the lysates, overnight at 4 °C, in a spinning rotator. On the following day, beads were centrifuged for 30 s at 2,300 g and 60 µL of supernatant was transferred to a microcentrifuge tube with 15 µL of 5 x SDS sample buffer and stored at -20 °C (flow-through sample for Western blot analysis). The rest of the supernatant was removed and beads were washed three times with IP buffer B, followed by centrifugation for 15 s at 3,300 g. After the last wash, all supernatant was removed and 75 µL of 2 x SDS sample buffer was added to the sample. IP samples were kept at -20 °C until Western blot analysis.

3.6 SDS–PAGE and Western blot

Samples were denatured at 95 °C for 20 min and loaded into a 10 % or 14 % polyacrylamide gel [40 % weight/volume (w/v) Acrylamide/Bis-acrylamide 37.5:1, BioRad]. NZYColour Protein

Marker II (NZYTech) was used as a protein ladder. Proteins were resolved by polyacrylamide gel electrophoresis (SDS–PAGE) at 150 V, for 1 h, and, then, transferred, for 2 h at 110 V, to polyvinylidene difluoride (PVDF) membranes, previously activated with methanol. Afterwards, membranes were blocked with 5 % (w/v) non-fat dry milk in 1 x tri-buffered saline (TBS) with 0.05 % (v/v) Tween 20 (Sigma-Aldrich), for 1 h. After that, membranes were incubated with the primary antibody, overnight at 4 °C, on a spinning rotator. On the next day, membranes were washed 3 times in 1 x TBS with 0.05 % (v/v) Tween 20 for 10 min and incubated with the secondary antibody for 1 h or 6 h, at room temperature. Then, membranes were washed again 3 times in 1 x TBS with 0.05 % (v/v) Tween 20 for 10 min.

Primary antibodies were diluted in 5 % (w/v) non-fat dry milk in 1 x TBS with 0.05 % (v/v) Tween 20 as follows: mouse anti-HDM2 4B2, mouse anti-HDM2 4B11[xi] and mouse anti-HDM2 SMP14 were diluted 1:500, 1:1,000 and 1:500, respectively, for HDM2 detection; rabbit anti-p53 CM1 (gift from Doctor Roman Hrstka, Masaryk Memorial Cancer Institute, Brno, Czech Republic) was diluted 1:10,000 for p53 detection; and mouse anti-α-tubulin (Sigma-Aldrich) was diluted 1:50,000 for α-tubulin detection. Secondary peroxidase-conjugated anti-mouse IgG (BioRad) or anti-rabbit IgG (BioRad) antibodies, both diluted 1:3,000 in 5 % (w/v) non-fat dry milk in 1 x TBS with 0.05 % (v/v) Tween 20 were used.

Antibody detection was performed incubating membranes with enhanced chemiluminescence reagent. The signal was imprinted in X-ray films, using different exposure times. The X-ray films were revealed in a Kodak Medical X-ray processor (Carestream Health).

To remove previously used antibodies and probe the membranes with new ones, we stripped off the membranes with NaOH. Briefly, dried membranes were re-activated with methanol, for 30 s and, then, hydrated in bidistilled water for 10 min. Afterwards, they were washed for 10 min in 1 x TBS with 0.05 % (v/v) Tween 20 and incubated with 250 mM NaOH for 10 to 15 min. This was followed by three 5-min washes in bidistilled water and three 5-min washes in 1 x TBS with 0.05 % (v/v) Tween 20. Then, membranes were blocked for 30 min in 5 % (w/v) non-fat dry milk in 1 x TBS with 0.05 % (v/v) Tween 20. After that, the previously described Western blot protocol was followed.

3.7 RNA extraction

Total RNA from transfected cells was extracted using Nucleospin® RNA II extraction kit (Macherey-Nagel), following manufacturer's instructions. An extra DNase treatment was performed, using DNase Q1 (Promega). RNA samples were then purified using the Nucleospin® RNA II extraction kit clean-up protocol.

3.8 RT–PCR analysis

cDNA was synthesized, according to manufacturer's instructions, either with NZY Reverse Transcriptase (NZYTech) or SuperScript II Reverse Transcriptase (Invitrogen), using either random hexamers (Invitrogen) and 1 µg of total RNA or Oligo $(dT)_{18}$ (NZYTech) and 1.5 µg of total RNA, respectively.

PCR was performed with GoTaq® Flexi Polymerase (Promega), according to standard procedures, using the resulting cDNA as template.

To check the integrity of the bicistronic transcripts, cDNAs obtained with NZY Reverse Transcriptase were PCR amplified with the following primers: Set I.A: 5'–GTCTCGAACTTAAGCTGCAG–3' (forward) and 5'–GCAAATCAGGTAGCCCAGG–3' (reverse); Set II.A: 5'–ATGGCTTCCAAGGTGTACGA–3' (forward) and 5'–ATCGATTTTACCACATTTGTAGAGG–3' (reverse); Set I.B: 5'–

[xi] Also a gift from Doctor Roman Hrstka, Masaryk Memorial Cancer Institute, Brno, Czech Republic

GGACGCTCCAGATGAAATGG–3' (forward) and 5'–CCGCCCCGACTCTAGAATTA–3' (reverse). *GAPDH* amplification was accomplished with 5'–ACCATCTTCCAGGAGCGAGAC–3' (forward) and 5'–GCCTTCTCCATGGTGGTGAA–3' (reverse). Different volumes of cDNA template (3 to 6 µL) at PCR solution were tested.

When cDNAs obtained with SuperScriptII were used, the integrity of the bicistronic transcripts was checked with: Set I.B: 5' – GGACGCTCCAGATGAAATGG – 3' (forward) and 5'–CCGCCCCGACTCTAGAATTA–3' (reverse); Set II.B: 5'–GTCTCGAACTTAAGCTGCAG–3' (forward) and 5'–GTGAGAGAAGCGCACACAG–3' (reverse). Several Mg^{2+} concentrations at PCR were tested (from 1 to 2 mM at PCR solution) as well as several annealing temperatures (from 55 to 58 °C).

3.9 Statistical analysis

Regarding bicistronic reporter constructs, RLuc is the internal control for transfection efficiency. Thus, cap-independent translation of FLuc was normalized to RLuc activity from the same construct. Then, for positive and negative controls (*PR_MYC_F* and *PR_HBB_F*, respectively), FLuc/RLuc ratio was normalized to that from the empty vector (*PR_F*) at each condition [0.1 % (v/v) DMSO, 1 µM Thapsigargin, 2 µM Thapsigargin or 4 µM Thapsigargin]. Afterwards, outliers were detected through interquartile range (IQR) analysis. Experiments with outliers in at least one of the controls were excluded. *PR_5'd160_F* FLuc relative expression was also normalized to that from the empty counterpart and IRQ analysis was also performed, followed by outliers' exclusion. Then, *PR_5'd160_F* FLuc relative expression was used as the normalizer to calculate variations in FLuc expression from *PR_d160_F*, *PR_5'd160R175H_F*, *PR_5'd160R248Q_F*, *PR_5'd160R273H_F* and *PR_5'd160R282W_F* bicistronic constructs. After normalizing FLuc/RLuc ratios from *PR_d160_F*, *PR_5'd160R175H_F*, *PR_5'd160R248Q_F*, *PR_5'd160R273H_F* and *PR_5'd160R282W_F* at 0.1 % (v/v) DMSO, 1 µM Thapsigargin, 2 µM Thapsigargin or 4 µM Thapsigargin to those from *PR_5'd160_F* at the same conditions, IQR analysis was, once again, performed. Outliers were, then, excluded for mean and standard deviation calculation.

In situations in which RLuc and FLuc activities had to be addressed separately (cryptic promoter activity assessment), either RLuc or FLuc activity was normalized to β-galactosidase activity (derived from the co-transfected β-galactosidase-containing plasmid, used as a control of transfection efficiency). RLuc/β-galactosidase or FLuc/β-galactosidase relative luciferase activities were, then, normalized to those from the empty counterpart (our negative control for cryptic promoter existence), to determine variations in RLuc and FLuc expression.

All results are expressed as mean ± standard deviation. Unpaired, two-tailed Student's test was used to estimate statistical significance, which was defined as * $P<0.05$, ** $P<0.01$ and *** $P<0.001$. All presented data with statistical significance result from at least three independent experiments.

4. Results and Discussion

4.1 Δ*160p53* IRES activity is repressed by its 5' UTR, but is partially recovered in the presence of the R175H p53 cancer mutation

Δ160p53 isoform expression was previously reported to be mediated by an IRES element located downstream of AUG160 start codon.[77] Since Δ160p53 was already implicated in tumor survival and proliferation[75], we aimed to understand the role of Δ160p53 IRES-mediated translation in carcinogenesis. Therefore, we proposed to evaluate the effect of Δ*160p53* 5' UTR, which was shown previously to have an inhibitory effect in Δ*160p53* IRES activity[77], on Δ160p53 IRES-mediated translation, as well as the effect of some p53 cancer mutations in counteracting Δ*160p53* 5' UTR inhibitory effect in this alternative mechanism of translation.

To achieve this goal, we used a bicistronic system based on Promega psiCHECK™–2 vector. This system has two reporter genes—*Renilla* Luciferase (RLuc) and firefly Luciferase (FLuc)—that are under the action of the same promoter. Therefore, transcription of both genes will produce a single bicistronic mRNA. While RLuc is the first cistron and its expression is driven by cap-dependent mechanisms, FLuc is the second cistron and is only translated if there is an upstream cloned sequence capable of promoting its translation through cap-independent mechanisms. RLuc and FLuc expression can be measured through bioluminescence assays since the reactions catalyzed by these enzymes emit light. RLuc catalyzes the oxidation of coelenterazine into coelenteramide. On the other hand, FLuc converts luciferin in oxyluciferin. First, we incubate samples with FLuc substrate-containing buffer and measure the bioluminescence produced by FLuc luciferase activity. Then, we incubate samples with RLuc substrate-containing buffer and measure the bioluminescence produced by RLuc luciferase activity. As the same buffer used to activate RLuc also induces quenching of FLuc bioluminescence, no FLuc bioluminescence signals are detected during RLuc bioluminescence measurements. Therefore, using a luminometer, it is possible to measure the amount of light produced in each reaction, which is directly proportional to the amount of protein expressed.[85] In our system, RLuc is an internal control and FLuc expression indicates the capacity of the sequence cloned upstream FLuc start codon to initiate translation through cap-independent mechanisms. To prevent translation reinitiation events after RLuc expression, a stable hairpin was previously cloned between the two cistrons, upstream of the sequence to be tested.[77] Furthermore, the presence of false positives regarding experimental limitations of this bicistronic system was evaluated and seems not to occur. These control experiments are explained in section 4.1.3.

4.1.1 Δ*160p53* 5' UTR represses Δ*160p53* IRES during endoplasmic reticulum stress conditions

Δ160p53 protein isoform is hardly detectable in cell lines expressing endogenous p53 in normal conditions.[77,75] However, ER stress was shown to induce expression of this shorter p53 isoform.[77] Moreover, it was also reported that induction of Δ160p53 is due to increased Δ160p53 expression and not due to decreased Δ160p53 degradation.[77] In the same study, Candeias and Marques-Ramos showed that Δ160p53 expression under ER stress induced by Thapsigargin[xii], known to impair cap-dependent

[xii] Thapsigargin inhibits ER Ca^{2+}-dependent ATPases, leading to Ca^{2+} depletion in ER, which, in turn, affects Ca^{2+}-dependent chaperones activity with a consequent increase of unfolded proteins.[151] ER unfolded proteins accumulation can be sensed by transmembrane sensors like PERK, whose activation leads to eIF2α phosphorylation and impairment of canonical translation through inhibition of ternary complex formation.[152]

translation, is accomplished by an IRES element located within Δ160p53 coding region.[77] They used two bicistronic constructs to limit Δ*160p53* IRES to the first 432 nucleotides of Δ160p53 coding region.[77] One of the bicistronic constructs used was that with RLuc and FLuc sequences already described in this thesis.[77] The other was a bicistronic construct with EGFP ORF at the first cistron and Δ160p53 sequence at the second cistron.[77] Moreover, cloning Δ133p53 at the second cistron, Candeias observed that the Δ*160p53* 5' UTR has an inhibitory effect on IRES activity.[77] This was observed in H1299 cells, a lung carcinoma p53-null cell line, and in A549 cells, a lung carcinoma cell line that expresses endogenous p53.[77] Δ*160p53* IRES activity was also stronger under ER stress conditions induced by 1 µM Thapsigargin, when compared to DMSO (vehicle) conditions.[77] Nevertheless, this was the first time that an IRES element was described for Δ160p53, pointing out the lack of information about this shorter p53 isoform.

Therefore, to see whether in HeLa cells—a different cell line from that used before—we could observe the same conclusions, thus strengthening the previously mentioned results, we performed bicistronic assays using the constructs represented in Figure 4.1. The empty vector *PR_F*, containing only RLuc and FLuc sequences, was used as a negative control for cap-independent mechanisms of translation. The *HBB* 5' UTR was previously cloned upstream FLuc AUG, originating the *PR_HBB_F* vector[83], the negative control for cap-independent translation initiation, since it cannot mediate IRES-dependent mechanisms of translation initiation. The positive control was the *PR_MYC_F* vector[83], containing *c-myc* IRES sequence, which has been a widely used positive control for IRES activity assessment.[86] Bicistronic constructs of Δ*160p53* IRES with or without its 5' UTR (*PR_5'd160_F* and *PR_d160_F*, respectively) were also available in the laboratory[77] and were used for evaluation of Δ*160p53* IRES activity in the presence or absence of Δ*160p53* 5' UTR.

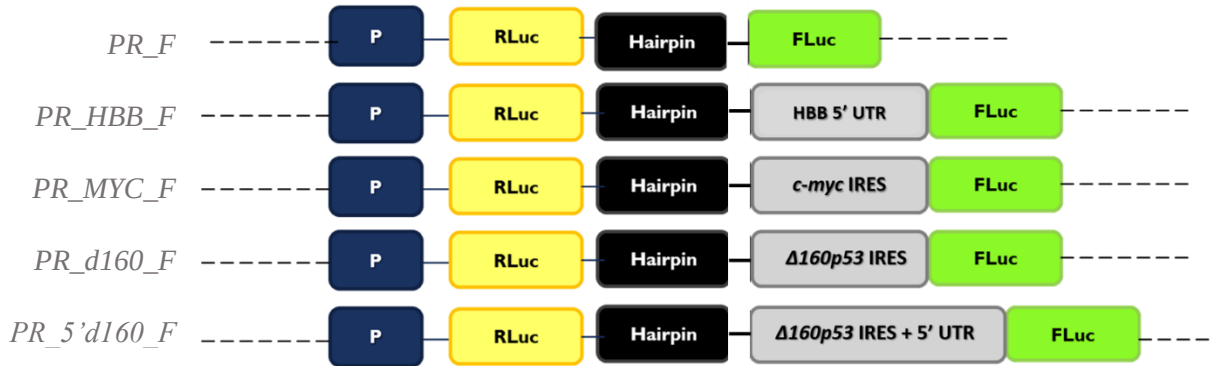


Figure 4. 1 — Schematic representation of the constructs used to address Δ*160p53* IRES activity in the presence or absence of Δ*160p53* 5' UTR in HeLa cells. RLuc is *Renilla* luciferase cap-dependent translated cistron (yellow box) and FLuc is firefly luciferase cap-independent translated cistron (green box). The blue box with the P letter represents the SV40 promoter. Grey boxes represent the cloned sequences upstream FLuc start codon. *PR_F* is the empty vector; *PR_HBB_F*, the *human β-globin* 5' UTR-containing construct and *PR_MYC_F*, the *c-myc* IRES-containing vector. *PR_d160_F* corresponds to Δ*160p53* IRES-containing vector, which contains the first 432 nucleotides of Δ160p53 coding sequence. *PR_5'd160_F* vector contains Δ*160p53* IRES (first 432 nucleotides of Δ160p53 coding sequence) and its 5' UTR (all Δ133p53 coding sequence up to Δ160p53 start codon, excluding Δ133p53 start codon). The black box represents the hairpin between the two cistrons.

HeLa cells, at 70–90 % cell confluence, were transfected with 375 ng of either *PR_F*, *PR_HBB_F*, *PR_MYC_F*, *PR_5'd160_F* or *PR_d160_F*. Afterwards, 24 h post-transfection, cells were treated with 1[77], 2 or 4 µM Thapsigargin or with the corresponding control vehicle, DMSO [0.1 % (v/v)]. Different Thapsigargin concentrations were tested in order to evaluate the presence of a dose-response and to select the Thapsigargin concentration that better induces Δ*160p53* IRES for further experiments (section 4.1.2). Sixteen hours later, cells were lysed and bioluminescence assays were performed. It is noteworthy that we can only compare FLuc expression from different constructs after excluding

transfection efficiency differences. In order to do that, we normalized FLuc bioluminescence signals to our internal control—RLuc. After obtaining FLuc/RLuc ratio for each transfected construct, we compared *PR_HBB_F* and *PR_MYC_F* FLuc/RLuc relative activity with that from the empty vector, for validation of the bicistronic system. An FLuc/RLuc relative activity greater than 1 indicates that the cloned sequence upstream FLuc is capable of driving cap-independent translation. This was the expected for *PR_MYC_F*. On the other hand, FLuc/RLuc relative activity ≤ 1, which is similar to that from the empty vector, indicates that the cloned sequence upstream FLuc cannot mediate a cap-independent initiation of translation. This was the expected for *PR_HBB_F*. Moreover, we also compared FLuc/RLuc relative activity of *PR_5'd160_F* (bicistronic construct containing Δ*160p53* IRES and its 5' UTR) with that from *PR_F* (Figure 4.2), to see how Δ*160p53* IRES behaved in the presence of its 5' UTR. We were expecting no IRES activity, as reported by Marques-Ramos (2013)[77]. Δ*160p53* IRES activity in the absence of its 5' UTR was assessed later by comparing FLuc/RLuc relative activity from *PR_d160_F* with that from *PR_5'd160_F*.

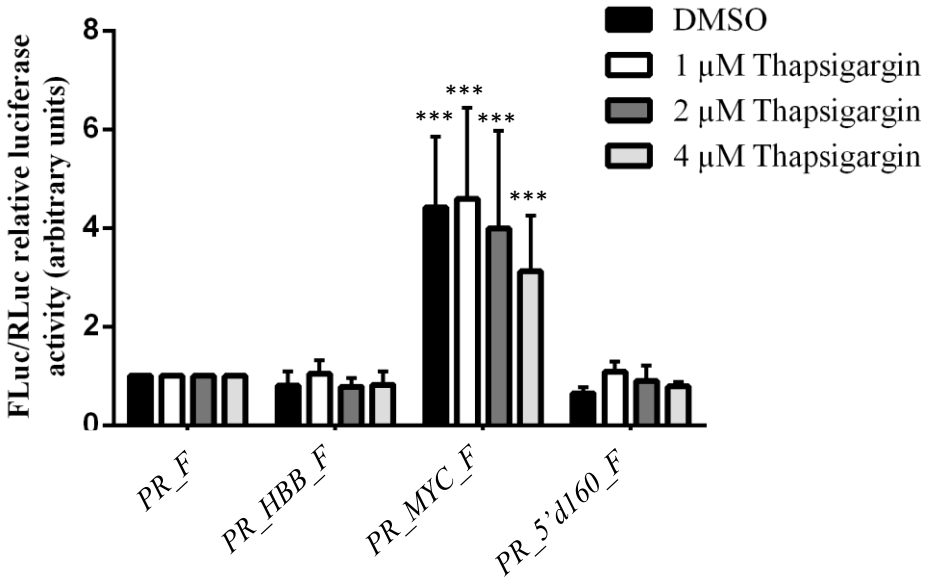


Figure 4. 2 — In the presence of its 5' UTR, Δ160p53 coding sequence cannot mediated cap-independent translation of FLuc, not even in endoplasmic reticulum stress conditions, in HeLa cells. HeLa cells were transfected with a bicistronic construct containing Δ160p53 ORF and its 5' UTR (*PR_5'd160_F*), or with one of the controls used in the experiment: the empty vector (*PR_F*), the *human β-globin* 5' UTR-containing construct (*PR_HBB_F*) and the *c-myc* IRES-containing vector (*PR_MYC_F*). Transfected cells were treated for 16 h with 1, 2 or 4 μM Thapsigargin or with DMSO (control vehicle). Presented data are the result of at least three independent experiments. Asterisks (*) indicate statistical significance in relation to the counterpart empty vector. *** P<0.001

In all tested Thapsigargin concentrations, FLuc/RLuc relative expression from *PR_5'd160_F* is similar to that from the empty vector *PR_F*. The relative Fluc/RLuc expression is 0.64-fold, 1.08-fold, 0.89-fold and 0.79-fold at DMSO, 1 μM Thapsigargin, 2 μM Thapsigargin and 4 μM Thapsigargin, respectively. These results suggest that Δ160p53 coding sequence with its 5' UTR is not able to drive cap-independent translation of Fluc in HeLa cells, in normal and in ER stress conditions. Regarding our negative and positive controls, Fluc/RLuc relative expression is what we expected. Accordingly, *PR_HBB_F* negative control cannot mediate cap-independent translation of FLuc, displaying 0.81-fold, 1.05-fold, 0.79-fold and 0.82-fold FLuc/RLuc relative expression at DMSO, 1 μM Thapsigargin, 2 μM Thapsigargin and 4 μM Thapsigargin, respectively, when compared to *PR_F* counterpart. Besides, we can observe 4.42-fold, 4.59-fold, 4.00-fold and 3.12-fold increase of *PR_MYC_F* FLuc/RLuc expression

levels at DMSO, 1 µM Thapsigargin, 2 µM Thapsigargin and 4 µM Thapsigargin, respectively, when compared to *PR_F*. These results confirm that *c-myc* cloned sequence can drive FLuc expression through a cap-independent manner. Moreover, we can see that FLuc/RLuc relative expression is similar in normal conditions (DMSO) and up to 2 µM Thapsigargin, with a greater cap-independent translation at 1 µM Thapsigargin. At 4 µM Thapsigargin, cap-independent translation of FLuc mediated by *c-myc* cloned sequence seems to be compromised.

After validating our system and confirming that positive and negative controls retrieve the results we expected, we can conclude that Δ160p53 coding sequence with its 5' UTR cannot efficiently mediate cap-independent translation of FLuc ORF in HeLa cells, in all tested conditions.

Then, the next step in this bioluminescence data analysis, is the comparison of *PR_d160_F* FLuc/RLuc relative expression to that from *PR_5'd160_F* (Figure 4.3), to assess the importance of Δ*160p53* 5' UTR on Δ*160p53* IRES activity.

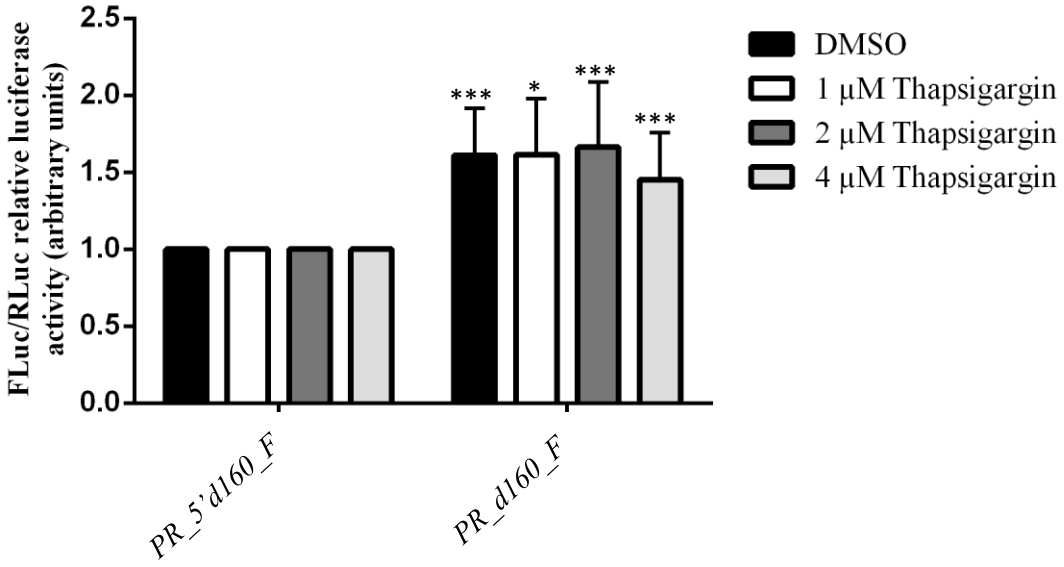


Figure 4. 3 — Δ*160p53* 5' UTR has an inhibitory effect on Δ*160p53* IRES activity, not only in normal, but also in endoplasmic reticulum stress conditions, in HeLa cells. HeLa cells were transfected with a bicistronic construct containing either Δ160p53 ORF (*PR_d160_F*) or Δ160p53 ORF and its 5' UTR (*PR_5'd160_F*). Transfected cells were treated for 16 h with 1, 2 or 4 µM Thapsigargin or with DMSO (control vehicle). Presented data are the result of at least three independent experiments. Asterisks (*) indicate statistical significance in relation to *PR_5'd160_F.* *P<0.05, *** P<0.001

In all tested Thapsigargin conditions, there is an increase in FLuc/RLuc relative expression from *PR_d160_F*, compared to that from *PR_5'd160_F*. It is observed a 1.61-fold, 1.61-fold, 1.66-fold and 1.45-fold increase at DMSO, 1 µM, 2 µM and 4 µM Thapsigargin, respectively. Therefore, in the absence of its 5' UTR, Δ160p53 coding sequence up to nucleotide 432 can drive cap-independent translation of FLuc, not only in normal but also in ER stress conditions. Furthermore, without its 5' UTR, ER stress induction by Thapsigargin does not stimulate Δ*160p53* IRES to a different extent than normal conditions, since FLuc/Rluc relative expression of *PR_160_F* is similar in normal and in Thapsigargin conditions up to 2 µM. These results are not the same obtained by Marques-Ramos, in 2013. In her experiments, ER stress leaded to a higher Δ*160p53* IRES induction then normal conditions.[77] These differences may be due to, for instance, the use of a different cell line than those from Marques-Ramos studies or loss Thapsigargin activity. One way to assess Thapsigargin activity could be to monitor total eIF2α and phosphorylated eIF2α by Western blot analysis, using the remain lysates of bioluminescence assays. Increased levels of phosphorylated eIF2α in conditions in which

Thapsigargin was used would indicate ER stress induction by Thapsigargin. Nevertheless, for 4 μM Thapsigargin, we see a different induction of Δ*160p53* IRES. With this Thapsigargin concentration, cap-independent translation of FLuc mediated by Δ*160p53* cloned sequence, although existing, seems to be compromised: only 1.45-fold increase *versus* more than 1.60-fold increase in normal conditions and in ER stress conditions up to 2 μM Thapsigargin. It is well known that while moderate ER stress triggers pro-survival responses, severe ER stress leads to activation of apoptotic pathways.[87] As Δ160p53 seems to promote cell survival and proliferation, its cap-independent expression may be only accomplished during normal and moderate ER stress conditions. Therefore, 4 μM Thapsigargin could be inducing severe ER stress, which, in turn, would be compromising Δ160p53 cap-independent expression. This is only a possible explanation for our results; however, we also observe a decrease in FLuc/RLuc relative expression in the positive control (*PR_MYC_F)* at 4 μM Thapsigargin, which strengthens our hypothesis. We may confirm this by Western blot analysis of pro-apoptotic factors, such as the C/EBP homologous protein (CHOP) and the apoptosis-signal-regulating kinase (ASK1).[87] Therefore, the differences between Marques-Ramos studies and our results may be explained through a biological perspective — different cell lines have different genes being expressed and consequently, different active pathways, which lead to different responses under the same stress conditions. Thus, in HeLa cells, unlike A549 and H1299 cells used by Marques-Ramos, ER stress may not induce Δ160p53 expression to a different extent than normal cellular conditions, and therefore this pathway may not be crucial for ER moderate stress response in this cell line.

All in all, we can conclude that Δ*160p53* 5' UTR seems to have an inhibitory effect on Δ*160p53* IRES activity. Since Δ160p53 expression was already associated with pro-oncogenic functions[75], it is reasonable that cells restrain this shorter p53 isoform expression. The mechanisms by which this 5' UTR inhibits Δ160p53 expression are not known yet but may be through disruption of Δ*160p53* IRES secondary structure or by interfering in ribosome and ITAFs interaction with *p53* mRNA. Nevertheless, the aim seems to be the cells protection from Δ160p53 pro-oncogenic traits. However, it is noteworthy that these assumptions bring an important question. If Δ160p53 has pro-oncogenic functions, why was this isoform preserved during evolution? Indeed, Δ160p53 is a conserved p53 isoform among mammals and was subjected to selective pressure for millions of years.[75] Thus, certain unfavorable conditions that disrupt cell homeostasis may induce Δ160p53 expression through an IRES-dependent manner, even in the presence of its 5' UTR. In those situations, Δ160p53 may confer adaptive advantage to the cell, explaining why this isoform was maintained during natural selection over the years. Which conditions trigger Δ160p53 expression are still not known, but, further investigation on the biological significance of Δ160p53 will be of great relevance to understand Δ160p53 crucial roles in cell homeostasis.

4.1.2 R175H p53 mutation can partially recover Δ*160p53* IRES activity in the presence of Δ*160p53* 5' UTR, under ER stress conditions

Mutations in *TP53* gene, which encodes for p53 tumor suppressor, are very common in cancer.[88] Research in this field has shown that these p53 mutations, besides abrogating FL-p53 tumor suppressor functions, also provide mutant p53 protein with new activities, named "gain-of-functions" (GOF), that can contribute to tumor survival and progression.[88] In 2016, Candeias *et al.* tested the effect of some p53 missense mutations, already associated with p53 tumorigenic activities (for instance R175H, R248Q and R273H GOF p53 mutations), in Δ160p53 expression.[75] In all cases, an increase in Δ160p53 levels was observed and was due to enhanced mRNA translation and not to increased protein stability,[75] indicating a role of GOF p53 mutations in Δ160p53 expression. Furthermore, Candeias *et al.* associated p53 tumorigenic functions, such as invasion and strong survival capacity, with Δ160p53 expression, which suggests that GOF phenotypes induced by p53 mutations depend on the shorter p53 isoforms.[75]

Thus, considering the aforementioned data, we hypothesized that p53 GOF mutations can induce Δ160p53 expression through its IRES element. Therefore, we proposed to study the effect of some common p53 GOF mutations on Δ*160p53* IRES activity. Our main goal was to see whether p53 mutations were able to counteract the inhibitory effect of Δ*160p53* 5' UTR on Δ160p53 cap-independent translation. We tested the three mutations studied by Candeias *et al.*[75]—R175H, R248Q and R273H—and another p53 missense mutation—R282W—, which had also been associated with p53 tumorigenic activities.[89] Bicistronic constructs containing Δ*160p53* IRES and its 5' UTR, and carrying R175H, R248Q and R273H mutations (*PR_5'd160R175H_F*, *PR_5'd160R248Q_F* and *PR_5'd160R273H_F*, respectively) were already available in the laboratory.[84] The bicistronic construct containing Δ*160p53* IRES and its 5' UTR, and carrying R282W mutation (*PR_5'd160R282W_F*) was obtained by site-directed mutagenesis, as described in Chapter 3.1, using *PR_5'd160_F* as template. All constructs are represented in Figure 4.4.

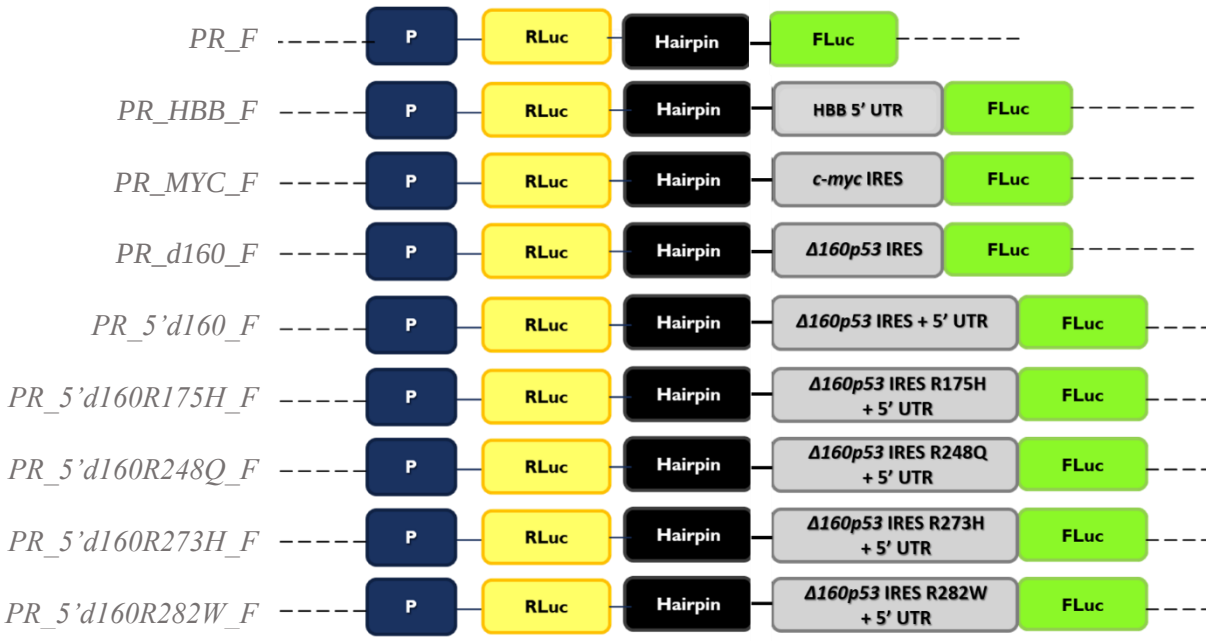


Figure 4. 4 — Schematic representation of the constructs used to address the impact of R175H, R248Q, R273H and R282W missense mutations on Δ*160p53* IRES activity in the presence of the Δ*160p53* 5' UTR. RLuc is *Renilla* luciferase cap-dependent translated cistron (yellow box) and FLuc is firefly luciferase cap-independent translated cistron (green box). The blue box with the P letter represents the SV40 promoter. Grey boxes represent the cloned sequences upstream FLuc start codon. *PR_F* is the empty vector; *PR_HBB_F*, the *human β-globin* 5' UTR-containing construct and *PR_MYC_F*, the *c-myc* IRES-containing vector. *PR_d160_F* corresponds to Δ*160p53* IRES-containing vector, which contains the first 432 nucleotides of Δ160p53 coding sequence. *PR_5'd160_F* vector contains Δ*160p53* IRES (first 432 nucleotides of Δ160p53 coding sequence) and its 5' UTR (all Δ133p53 coding sequence up to Δ160p53 start codon, excluding Δ133p53 start codon). *PR_5'd160R175H_F*, *PR_5'd160R248Q_F*, *PR_5'd160R273H_F* and *PR_5'd160R282W_F* are the bicistronic constructs containing Δ*160p53* IRES and its 5' UTR, and carrying R175H, R248Q, R273H and R282W p53 mutations, respectively. The black box represents the hairpin between the two cistrons.

HeLa cells, at 85–90 % cell confluence, were then transfected with 375 ng of either *PR_F*, *PR_HBB_F*, *PR_MYC_F*, *PR_5'd160_F*, *PR_d160_F*, *PR_5'd160R175H_F*, *PR_5'd160R248Q_F*, *PR_5'd160R273H_F* or *PR_5'd160R282W*. Twenty-four hours post-transfection, cells were treated with either the control vehicle, DMSO [0.1 % (v/v)], or 2 μM[xiii] Thapsigargin. Sixteen hours later, cells were lysed, and bioluminescence assays were performed. Since cell over-confluence also seems to induce Δ*160p53* IRES activity[77], in this assay we decided to transfect at a higher cell confluence than

[xiii] Thapsigargin concentration that allowed a better induction of Δ*160p53* IRES in section 4.1.1.

before (85–90 % cell confluence *versus* 70–90 % cell confluence) to create over-confluence conditions at the time of drug treatments.

As done before, we validated our experimental results by normalizing *PR_HBB_F* and *PR_MYC_F* FLuc/RLuc relative expression levels to that from *PR_F*. At the same time, we also compared *PR_5'd160_F* FLuc/RLuc ratio to that from the empty counterpart (Figure 4.5).

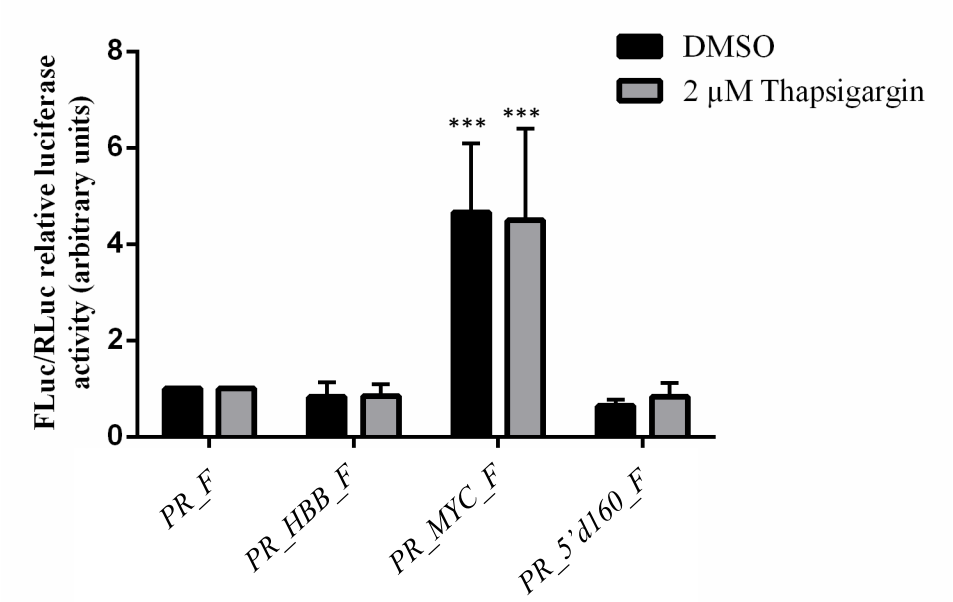


Figure 4. 5 — In HeLa cells transfected at 80-90 % cell confluence, Δ*160p53* IRES in the presence of its 5' UTR cannot mediate cap-independent translation of FLuc. HeLa cells were transfected with a bicistronic construct containing Δ160p53 ORF and its 5' UTR (*PR_5'd160_F*), or with one of the controls used in the experiment: the empty vector (*PR_F*), the *human β-globin* 5' UTR-containing construct (*PR_HBB_F*) and the *c-myc* IRES-containing vector (*PR_MYC_F*). Transfected cells were treated for 16 h with either 2 µM Thapsigargin or DMSO (control vehicle). Presented data are the result of at least three independent experiments. Asterisks (*) indicate statistical significance in relation to the counterpart empty vector. *** P<0.001

As before, Δ*160p53* IRES, in the presence of its 5' UTR, is not able to mediate FLuc expression through a cap-independent manner. Furthermore, combining ER stress, induced by 2 µM Thapsigargin, with cellular stress, induced by cell over-confluence, is not enough to activate Δ*160p53* IRES and overcome the inhibitory effect of its 5' UTR, in HeLa cells. In fact, FLuc/RLuc relative expression of *PR_5'd160_F* is similar to that from *PR_F* (0.64-fold and 0.83-fold at DMSO and at 2 µM Thapsigargin, respectively). FLuc/RLuc relative expression from *PR_HBB_F* negative control is what we expected, with 0.83-fold and 0.84-fold at DMSO and at 2 µM Thapsigargin, respectively, compared to the empty vector. Once again, *PR_MYC_F* positive control is able to drive FLuc cap-independent translation. We observe a 4.67-fold and a 4.50-fold increase in FLuc/RLuc relative expression, in relation to *PR_F*, at DMSO and at 2 µM Thapsigargin, respectively.

After validating the experiment through bioluminescence analysis of positive and negative controls (*PR_MYC_F* and *PR_HBB_F*, respectively), we compared the effect of p53 cancer mutations on inducing Δ160p53 cap-independent expression, in the presence of Δ160p53 5' UTR. In order to do that, we compared FLuc/RLuc relative expression of *PR_d160_F*, *PR_5'd160R175H_F*, *PR_5'd160R248Q_F, PR_5'd160R273H_F* and *PR_5'd160R282W_F* to that from *PR_5'd160_F* (Figure 4.6).

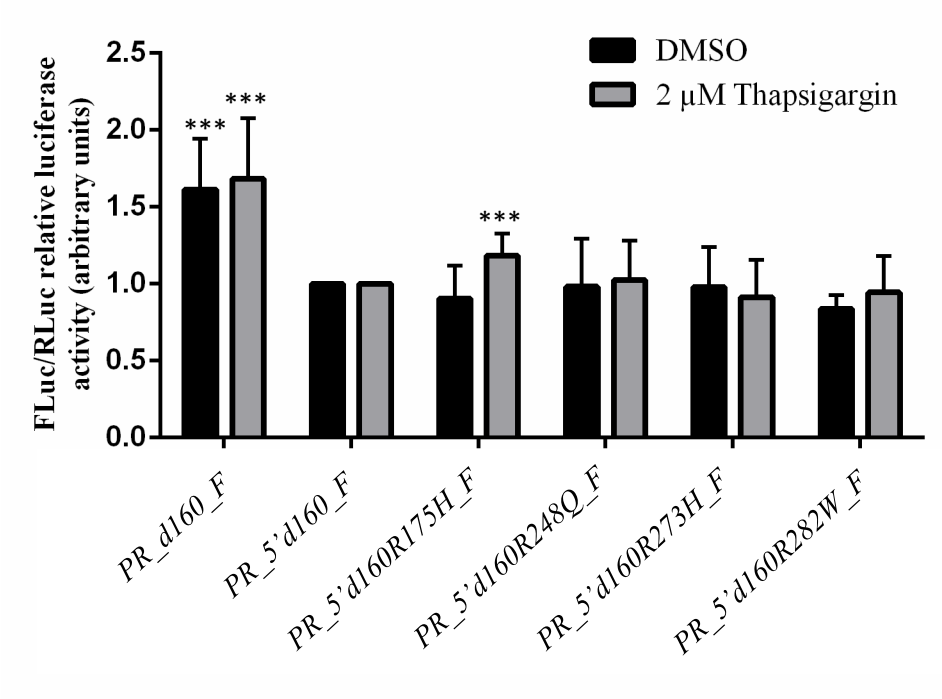


Figure 4. 6 — p53 missense mutation R175H can revert some of the inhibitory effect of Δ*160p53* 5' UTR on Δ*160p53* IRES activity, in ER stress conditions. HeLa cells were transfected with a bicistronic construct containing either Δ160p53 ORF (*PR_d160_F*) or Δ160p53 ORF and its 5' UTR without (*PR_5'd160_F*) and with p53 missense mutations (*PR_5'd160R175H_F, PR_5'd160R248Q_F, PR_5'd160R273H_F* and *PR_5'd160R282W_F*). Transfected cells were treated for 16 h with either 2 μM Thapsigargin or DMSO (control vehicle). Presented data are the result of at least three independent experiments. Asterisks (*) indicate statistical significance in relation to *PR_5'd160_F*. *** P<0.001

In the tested conditions, FLuc/RLuc relative expression from *PR_d160_F* is 1.61-fold and 1.69-fold higher, at DMSO and at 2 μM Thapsigargin, respectively, than that from *PR_5'd160_F*. This confirms the inhibitory effect of Δ*160p53* 5' UTR on Δ*160p53* IRES activity, since without its 5' UTR, Δ*160p53* can mediate cap-independent translation of FLuc. As for relative FLuc/RLuc expression levels from *PR_5'd160R248Q_F*, *PR_5'd160R273H_F* and *PR_5'd160R282W_F*, they are similar to that from *PR_5'd160_F*—0.98-fold versus 1.02-fold, 0.98-fold versus 0.91-fold, 0.83-fold versus 0.94-fold, respectively, at DMSO and 2 μM Thapsigargin. For *PR_5'd160R175H_F* bicistronic construct, FLuc/RLuc relative expression is similar to that from *PR_5'd160_F* at DMSO (0.81-fold); however, at 2 μM Thapsigargin, FLuc/RLuc relative expression is 1.18-fold higher than that from *PR_5'd160_F*. Therefore, R175H p53 mutation can counteract some of the inhibitory effect of Δ*160p53* 5' UTR on Δ*160p53* IRES activity, under ER stress conditions, in HeLa cells. However, this mutation cannot restore all Δ*160p53* IRES activity, since FLuc/RLuc relative expression from *PR_5'd160R175H_F* is lower than that observed from *PR_d160_F*, at 2 μM Thapsigargin (1.18-fold versus 1.69-fold, compared to *PR_5'd160_F*).

It is noteworthy that all tested mutations are within p53's DNA-binding domain[90], and hence, they can alter p53 function as a transcription factor.[91] However, Δ160p53 lacks part of the DNA-binding domain.[69] Thus, although playing an important role in protein function of p53 isoforms that contain an integral DNA-binding domain, the studied p53 mutations may not have impact in Δ160p53 protein function, and may act through different mechanisms in Δ160p53 regulation. Indeed, Candeias *et al.* showed that R273H mutation do not affect Δ160p53 protein functions.[75] Furthermore, R175H, R248Q and R273H p53 mutations have already shown to increase Δ160p53 protein levels, not by inhibiting Δ160p53 degradation but by inducing its translation.[75] Being in agreement with this, we showed in this thesis that R175H mutation can revert some of the inhibitory effect of Δ*160p53* 5' UTR on Δ*160p53*

IRES activity, during ER stress, and therefore, induce FLuc cap-independent translation. Still, although being associated with Δ160p53 overexpression[75], R248Q and R273H missense mutations do not show any role in cap-independent expression of FLuc, in HeLa cells. Candeias *et al*. (2016)[75] used a different system and a different experimental set up to address the role of R248Q and R273H missense mutations on Δ160p53, so, our results may differ from theirs. For instance, they used different cells lines. In fact, each cell line is unique and presents differences in gene expression patterns. Hence, auxiliary proteins involved in IRES-dependent expression of Δ160p53 may not be expressed in all cancer cell lines. As oncogenic cells, HeLa cells may have mechanisms to survive and proliferate that do not depend on Δ160p53 expression. Moreover, if R175H mutation is the most efficient of all tested mutations in counteracting Δ*160p53* 5' UTR, we may be able to see its effect in FLuc cap-independent expression in HeLa cells, but not the effect of R248Q, R273H and R282W missense mutations. Therefore, testing endogenous Δ160p53 levels in HeLa cells, in normal and in ER stress conditions, might give us insight about the role of Δ160p53 in this cell line survival. Furthermore, selection of cell lines expressing endogenous Δ160p53, through Western blot analysis of endogenous Δ160p53 levels, followed by bicistronic assays for Δ*160p53* IRES activity assessment could also be an interesting approach to follow. Indeed, to fully understand the role of p53 GOF mutations in Δ160p53 IRES-mediated translation and tumorigenesis, we need to use cell lines with active mechanisms to mediate Δ160p53 cap-independent expression. Furthermore, the identification of cancer cell lines in which Δ160p53 IRES-mediated expression is crucial for cell survival and oncogenic traits may open doors to a new cancer therapy targeting *Δ160p53* IRES. Additionally, other stress conditions, such as DNA damage, may induce Δ*160p53* IRES to a better extent. For instance, evaluation of Δ160p53 IRES-mediated translation after etoposide or UV irradiation exposure could also be performed, to complete our results and unveil new understanding of Δ160p53 mutant functions in cancer. At the end, a combined role of p53 mutations in *Δ160p53* IRES induction may also exist. Although it is very rare to have two missense mutations in *TP53* in cancer at the same time, a combined role of one missense mutation with synonymous mutations may happen.[92] The presence of more than one p53 mutation in *p53* transcripts may remodel mRNA secondary structure and stabilize *Δ160p53* IRES, thus favoring Δ160p53 IRES-mediated translation. Further studies involving bicistronic constructs with some combinations of missense and synonymous p53 mutations can be used to confirm this hypothesis.

4.1.3 FLuc expression from all tested bicistronic constructs does not seem to be a consequence of either alternative splicing or cryptic promoter activity

In order to understand the biological implications of the obtained results from sections 4.1.1 and 4.1.2, experimental limitations should be considered. Indeed, bicistronic construct containing RLuc and FLuc ORFs may produce false positives. For instance, we observed FLuc expression from empty vector and from *PR_HBB_F* negative control, which should not happen, since these constructs cannot mediate non-canonical translation initiation. Moreover, FLuc expression was a thousand times higher than that from non-transfected samples (data not shown), thus this was not due to background and cell autofluorescence. When using a bicistronic construct, reinitiation, read-through or ribosome shunting must be avoided, since they can promote FLuc IRES-independent translation. As a stable hairpin was previously cloned between RLuc and FLuc ORFs, reinitiation of canonical translation at FLuc ORF is unlikely. The same occurs for read-through, since a second stop codon was added after RLuc stop codon.[77] Thus, FLuc bioluminescence signals from empty vector and negative control could be due to ribosome shunting, according to which the 40S ribosomal subunit can shunt across an upstream AUG or stable hairpin structure by means of interactions between the downstream ORF that is about to be translated and the ribosomal RNA (rRNA) of the 40S subunit.[93] Since this mechanism involves base-

pairing between the rRNA of the 40S ribosomal subunit and the mRNA[93], a sequence alignment, followed by complementarity analysis could be done to exclude or not this non-canonical mechanism of translation as the cause for FLuc bioluminescence signals from empty vector and negative control. Nevertheless, we took in consideration the differences in transfection efficiency of different constructs and we did not compare absolute values of FLuc expression. In fact, we normalized FLuc expression of each construct to the correspondent RLuc expression. Therefore, if ribosome shunting did occur, all constructs should have been affected similarly and this effect should have been abolished. Notwithstanding, in the presence of weak IRES, which is common in cellular IRES, this effect might mask their activity. For that reason, site-directed mutagenesis of FLuc AUG start codon could be a good option to avoid FLuc independent translation through ribosome shunting, allowing more reliable results.

Besides this issue, cloning the sequences of interest in bicistronic constructs may create cryptic promoters or trigger aberrant alternative splicing events, which can lead to false-positive results.[94] These two questions are addressed in the following paragraphs.

4.1.3.1 FLuc expression from all tested bicistronic constructs does not seem to be a consequence of alternative splicing

Evaluating the existence of an IRES element in an mRNA requires the performance of control experiments to ensure that positive results are not in fact false positives. Cloning a new DNA sequence in a vector can lead to the appearance of new splice sites in the corresponding mRNA molecule. This can trigger aberrant splicing events that may originate a monocistronic transcript encoding only FLuc, which in turn will be translated through the canonical cap-dependent mechanism. This would increase the levels of FLuc protein measured by luminometry assays, since they would be the result of cap-dependent translation from the monocistronic transcript plus possible IRES-mediated translation mediated by Δ160p53 cloned sequences.

Therefore, to exclude alternative splicing events, HeLa cells were transfected with 3.75 µg of either *PR_F*, *PR_HBB*, *PR_MYC*, *PR_d160_F*, *PR_5'd160_F*, *PR_5'd160R175H_F*, *PR_5'd160R248Q_F*, *PR_5'd160R273H_F* or *PR_5'd160R282W_F*. Then, RNA was extracted from cells and the integrity of the bicistronic mRNAs transcribed from the equivalent transfected plasmids DNA was analyzed by RT–PCR, using two different sets of primers: Set I, which covers all RLuc ORF, the cloned sequence and part of FLuc ORF; and Set II, which covers part of RLuc ORF, the cloned sequence and all FLuc ORF (Figure 4.7). Experimental optimizations, which will be briefly described below, were necessary for mRNA integrity analysis.

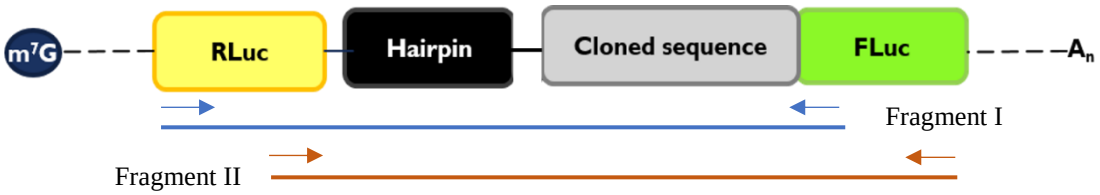


Figure 4. 7 — Schematic representation of a putative bicistronic mRNA transcribed from the equivalent transfected plasmid DNA. The blue circle with m^7G represents the 7-methylguanosine cap structure. RLuc is Renilla luciferase cap-dependent translated cistron (yellow box) and FLuc is firefly luciferase cap-independent translated cistron (green box). The black box represents the hairpin between the two cistrons and the grey box represents the cloned sequence upstream FLuc start codon. The poly(A) tail is represented by An. Arrows indicate the primers for complementary DNA (cDNA) amplification, in order to obtain fragments I (blue line) and II (brown line), which in turn will be resolved in an agarose gel, to provide insight on RNA integrity.

The first step in alternative splicing analysis by RT–PCR is complementary DNA (cDNA) synthesis of total extracted RNA. To do that, we started by using NZYReverse transcriptase, random

hexamers and 1 µg of total RNA as template. However, when performing PCR reaction, after cDNA synthesis, we could not observe any amplification. We tried different sets of primers for fragment I and fragment II amplification, different volumes of cDNA template (3 to 6 µL of the 20 µL reverse transcription reaction) and different number of cycles on PCR program (35 to 40 cycles). Nevertheless, the outcome was always the same: no cDNA amplification. One example is presented at Figure 4.8.

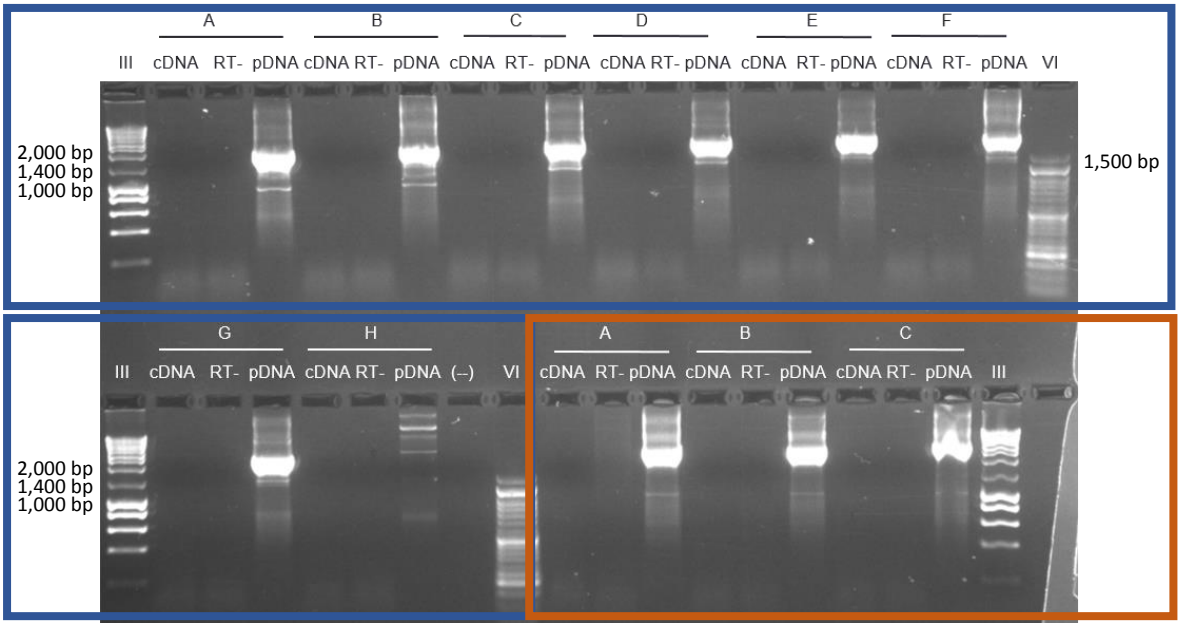


Figure 4. 8 — RT–PCR of bicistronic mRNAs produces no cDNA amplification. Agarose gel only shows positive controls (plasmid DNA, pDNA) amplification. A, B, C, D, E, F, G and H are, respectively, the RT–PCR of *PR_F*, *PR_HBB_F*, *PR_MYC_F*, *PR_d160_F*, *PR_5'd160_F*, *PR_5'd160R175H_F*, *PR_5'd160R248Q_F* and *PR_5'd160R273H_F* bicistronic mRNAs. III and VI are NZYLadderIII and NZYLadderVI DNA molecular weight ladders (NZYTech); RT- indicates the PCR amplification reaction without cDNA synthesis step, proving no DNA contamination occurred in the cDNA sample. The (-) lane contains the PCR negative control. PCR reactions with primers' set I.A (for fragment I amplification) are represented within the blue frames—pDNA fragments size: 2,105 base pairs (bp) (A), 2,155 bp (B), 2,455 bp (C), 2,534 bp (D) and 2,614 bp (E, F, G and H)—; within the orange frame are represented the PCR reactions with primers' set II.A (for fragment II amplification)—pDNA fragments size: 2,998 bp (A), 3,048 bp (B) and 3,348 bp (C). RT–PCR conditions: 3 µL of cDNA synthesis reaction; 2 mM of Mg^{2+} at final PCR reaction; PCR program: 95 ˚C for 3 min (initial denaturation), [95 ˚C for 30 s (denaturation), 55 ˚C for 30s (annealing), 72 ˚C for 1 min 30 s (extension)] x 35, 72 ˚C for 10 min (final extension).

To ensure that NZYreverse transcriptase is functional and promoting cDNA synthesis, thus excluding this step as the cause of no cDNA amplification, we performed RT–PCR for detection of an endogenous transcribed gene. We selected the housekeeping *GAPDH* gene as our control, since it is stably and constitutively expressed at high levels in many tissues and cells, including in HeLa cells.[95] RT–PCR analysis revealed amplification of a 100 bp fragment, which confirms that NZYreverse transcriptase preserves its enzymatic activity (Figure 4.9).

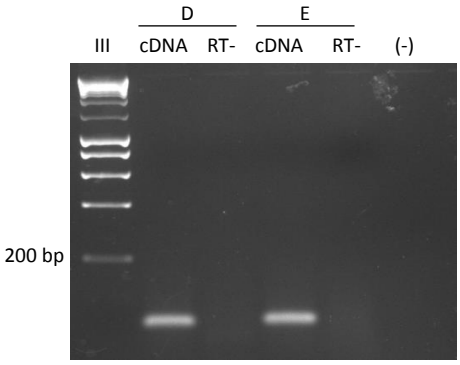


Figure 4. 9 — RT–PCR of *GPDH* mRNA reveals that NZYReverse transcriptase is functional. D and E are, respectively, the RT–PCR analysis for *GAPDH* mRNA from cells transfected with *PR_d160_F* and *PR_5'd160_F* plasmids. III corresponds to NZYLadderIII DNA molecular weight ladder (NZYTech); RT- indicates the PCR amplification reaction without cDNA synthesis step, proving no DNA contamination occurred in the cDNA sample. The (-) lane contains the PCR negative control. Fragment size: 100 base pairs (bp). RT–PCR conditions: 3 µL of cDNA synthesis reaction; 2 mM of Mg^{2+} at final PCR reaction; PCR program: 95 ˚C for 3 min (initial denaturation), [95 ˚C for 30 s (denaturation), 55 ˚C for 30s (annealing), 72 ˚C for 1 min 30 s (extension)] x 35, 72 ˚C for 10 min (final extension).

Then, we continued with PCR optimizations, testing other sets of primers, in combination with different cDNA volumes and different PCR cycles. As we can see in figure 4.10, we were able to obtain fragment I amplification with primer's set I.B from *P_5'd160R273H_F* cDNA. However, the control reaction without reverse transcriptase (RT-) also shows amplification (Figure 4.10).

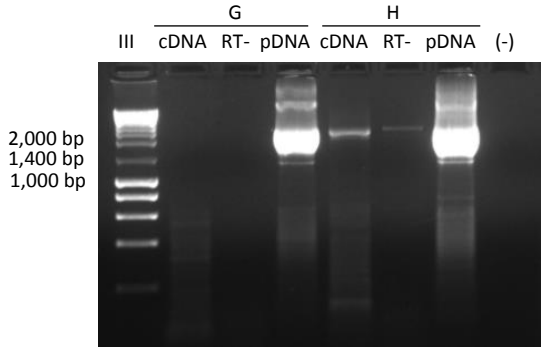


Figure 4. 10 — RT–PCR assay using *PR_5'd160R273H_F* mRNA shows cDNA amplification. G and H are, respectively, the RT–PCR of *PR_5'd160R248Q_F* and *PR_5'd160R273H_F* bicistronic mRNAs. III corresponds to NZYLadderIII DNA molecular weight ladder (NZYTech); RT- indicates the PCR amplification reaction without cDNA synthesis step, proving DNA contamination occurred in the cDNA sample. pDNA indicate the plasmid DNA positive control. The (-) lane contains the PCR negative control. Fragments size: 2,503 base pairs (bp) (RT- of H and pDNA of G and H), 2,370 bp (cDNA of G and H). Plasmid DNA fragment is bigger than the corresponding cDNA because it includes a chimeric intron that is removed during mRNA processing. RT–PCR conditions: 5 µL of cDNA synthesis reaction; 2 mM of Mg^{2+} at final PCR reaction; PCR program: 95 ˚C for 3 min (initial denaturation), [95 ˚C for 30 s (denaturation), 55 ˚C for 30s (annealing), 72 ˚C for 1 min 30 s (extension)] x 35, 72 ˚C for 10 min (final extension). PCR reaction was performed with primer's set I.B.

Given that the control reaction without reverse transcriptase (RT-) shows DNA amplification, indicating that there is DNA contamination, probably by inefficient DNA removal and degradation during RNA extraction step involving DNaseI incubation, we performed an extra DNase treatment to RNA samples, before cDNA synthesis. However, after this new step and after reproducing the PCR conditions described in Figure 4.10, we were not able to obtain cDNA amplification. Once again, we questioned NZYReverse transcriptase enzymatic activity. Information about this enzyme claims that it can be used for cDNA synthesis of up to 7 kb mRNAs, which, in turn, can be in low concentrations. In

fact NZYReverse transcriptase can synthesize cDNA of only 10 pg of template RNA.[96] Nevertheless, if NZYReverse transcriptase had lost some of its enzymatic activity, it could be able to synthesize cDNA of short and high concentrated transcripts, such as GAPDH (Figure 4.9), but not be able to synthesize long cDNA that are in very low concentration in the cell, as the bicistronic studied in this work. Therefore, we tested another reverse transcriptase—Superscript II (Invitrogen)—for cDNA synthesis. This enzyme is also an engineered reverse transcriptase that can produce cDNA of templates containing up to 12.3 kb[97],but the results were the same as before (data not shown). Maintaining Superscript II reverse transcriptase, we changed the use of random hexamers to Oligo(dT)$_{18}$ for cDNA production. While random hexamers allow the cDNA synthesis of all RNA species, since they have a random binding (no template specificity), with Oligo(dT), only mRNA will be used as template for cDNA production. As we want to evaluate the integrity of mRNAs, Oligo(dT) can be used. With Oligo(dT), we are reducing the amount of cDNA molecules that are produced, which in turn might improve PCR results, since high cDNA concentrations may difficult polymerase diffusion and action.[98] In fact, this last change proved to be the key for cDNA amplification (Figure 4.11). Nevertheless, similar unspecific bands are detected in the cDNA, RT- and also in the pDNA lane where purified plasmid DNA was used directly on the PCR.

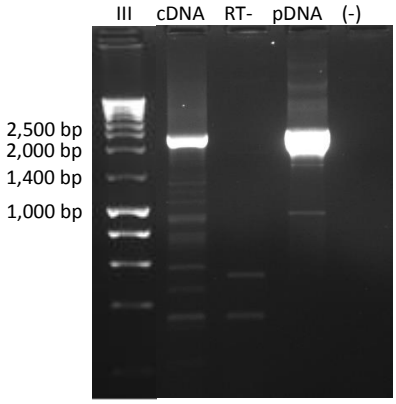


Figure 4. 11 — RT–PCR analysis using *PR_HBB_F* mRNA shows cDNA amplification. Besides 1,805 base pairs (bp) fragment amplification on cDNA lane, we also see some unspecific bands. RT- lane also has two fragments with low molecular weight. III corresponds to NZYLadderIII DNA molecular weight ladder (NZYTech); RT- indicates the PCR amplification reaction without cDNA synthesis step, proving DNA contamination occurred in the cDNA sample. The (-) lane contains the PCR negative control. Fragment size of pDNA: 1,938 bp. Plasmid DNA fragment is bigger than the corresponding cDNA, because it includes a chimeric intron that is removed during mRNA processing. RT–PCR conditions: 4 µL of cDNA synthesis reaction; 2 mM of Mg^{2+} at final PCR reaction; PCR program: 95 ˚C for 3 min (initial denaturation), [95 ˚C for 30 s (denaturation), 56 ˚C for 30s (annealing), 72 ˚C for 2 min (extension)] x 40, 72 ˚C for 10 min (final extension). PCR reaction was performed with primer's set I.B.

To eliminate unspecific bands, we started by fixing the amount of cDNA used as template in PCR reaction. From the cDNA synthesized with 1.5 µg of total RNA in a reaction volume of 20 µL, we used 3 µL as template for PCR. Then, for primers' set I.B, augmenting the annealing temperature and reducing Mg^{2+} levels in PCR reaction allowed us the removal of unspecific amplifications. Indeed, it is known that augmenting the annealing temperature reduces unspecific primer binding and increases annealing stringency.[99] Besides, excessive Mg^{2+} concentrations can reduce polymerase fidelity and increase the levels of unspecific amplifications.[100–102] Fragment I amplification of all tested mRNAs, with the exception for those from transfected *PR_5'd160R175H_F* and *PR_5'd160R248Q_F* was, then, accomplished with 1.125 mM of Mg^{2+} at PCR reaction, using 58 ˚C as the annealing temperature and performing 37 PCR cycles. *PR_5'd160R175H_F* and *PR_5'd160R248Q_F* transcribed mRNAs required 1.5 mM Mg^{2+} at PCR reaction (Figure 4.12).

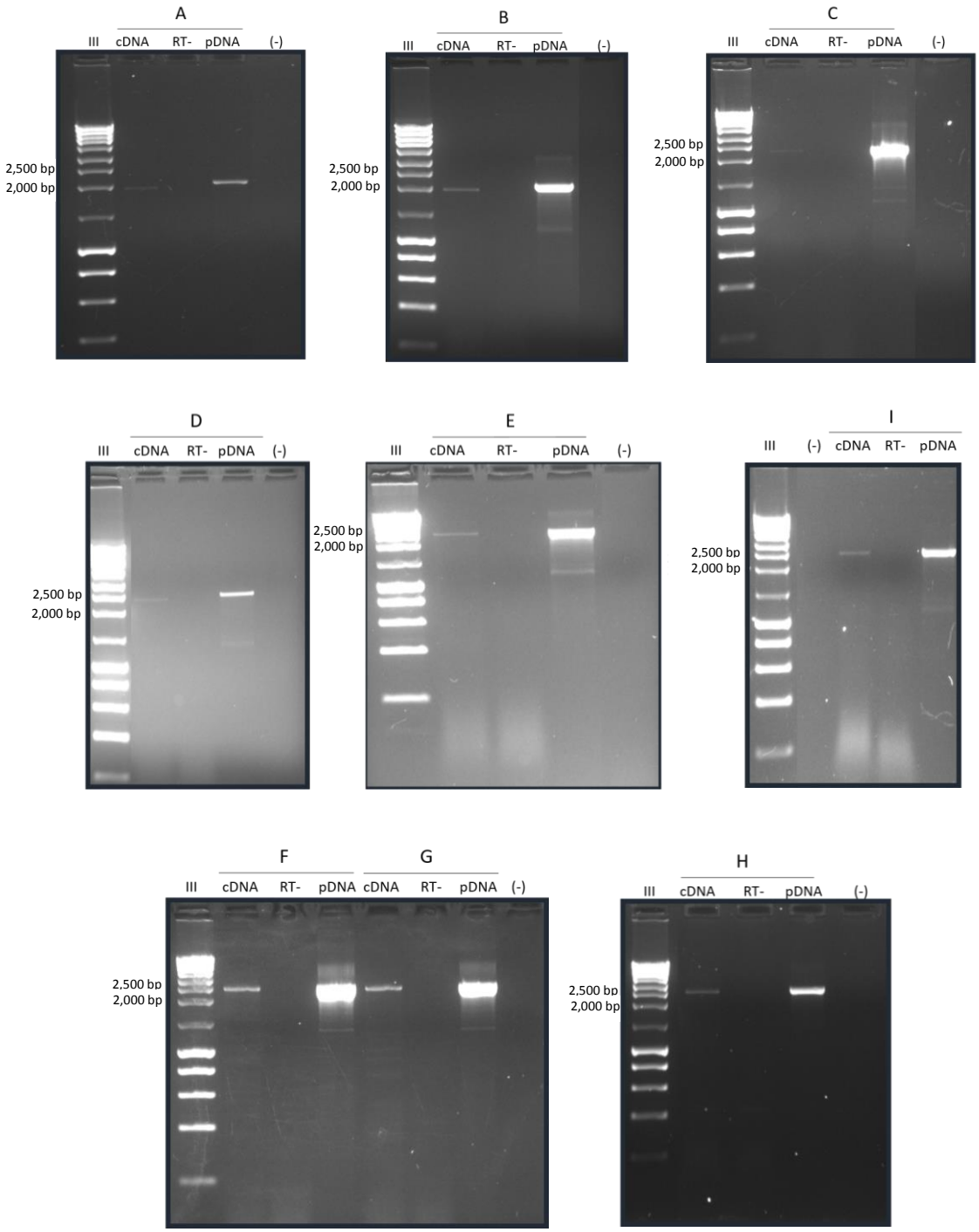


Figure 4. 12 — Analysis of the bicistronic mRNAs by RT–PCR using primers' set I.B reveals the correct full-length DNA fragment, confirming the full-length expressed mRNA. A, B, C, D, E, F, G, H and I are, respectively, the RT–of *PR_F*, *PR_HBB_F*, *PR_MYC_F*, *PR_d160_F*, *PR_5'd160_F*, *PR_5'd160R175H_F*, *PR_5'd160R248Q_F*, *PR_5'd160R273H_F* and *PR_5'd160R282W_*F bicistronic mRNAs. III corresponds to NZYLadderIII DNA molecular weight ladder (NZYTech); RT- indicates the PCR amplification reaction without cDNA synthesis step, proving no DNA contamination occurred in the cDNA sample. The (-) lanes contain the PCR negative controls. cDNA fragment sizes: 1,755 base pairs (bp) (A), 1,805 bp (B), 2,105 bp (C), 2,290 bp (D) and 2,370 bp (E, F, G, H and I). pDNA fragments are 133 bp bigger than the corresponding cDNA because it includes a chimeric intron that is removed during mRNA processing. RT–PCR conditions for A, B, C, D, E, H and I: 3 µL of cDNA synthesis reaction; 1.125 mM of Mg^{2+} at final PCR reaction; PCR program: 95 ˚C for 3 min (initial denaturation), [95 ˚C for 30 s (denaturation), 58 ˚C for 30s (annealing), 72 ˚C for 2 min (extension)] x 37, 72 ˚C for 10 min (final extension). RT–PCR conditions for F and G: 3 µL of cDNA synthesis reaction; 1.5 mM of Mg^{2+} at final PCR reaction; PCR program: 95 ˚C for 3 min (initial denaturation), [95 ˚C for 30 s (denaturation), 58 ˚C for 30s (annealing), 72 ˚C for 2 min (extension)] x 37, 72 ˚C for 10 min (final extension).

After obtaining fragment I amplification for all expressed bicistronic mRNAs, we moved onto fragment II RT–PCR optimizations. We started by maintaining the PCR conditions used before, and then, optimize to remove unspecific bands. Primers' set II.B was used in a PCR reaction with 1 mM Mg^{2+}, 57 ˚C as the annealing temperature and 35 PCR cycles. The obtained fragments are depicted in Figure 4.13.

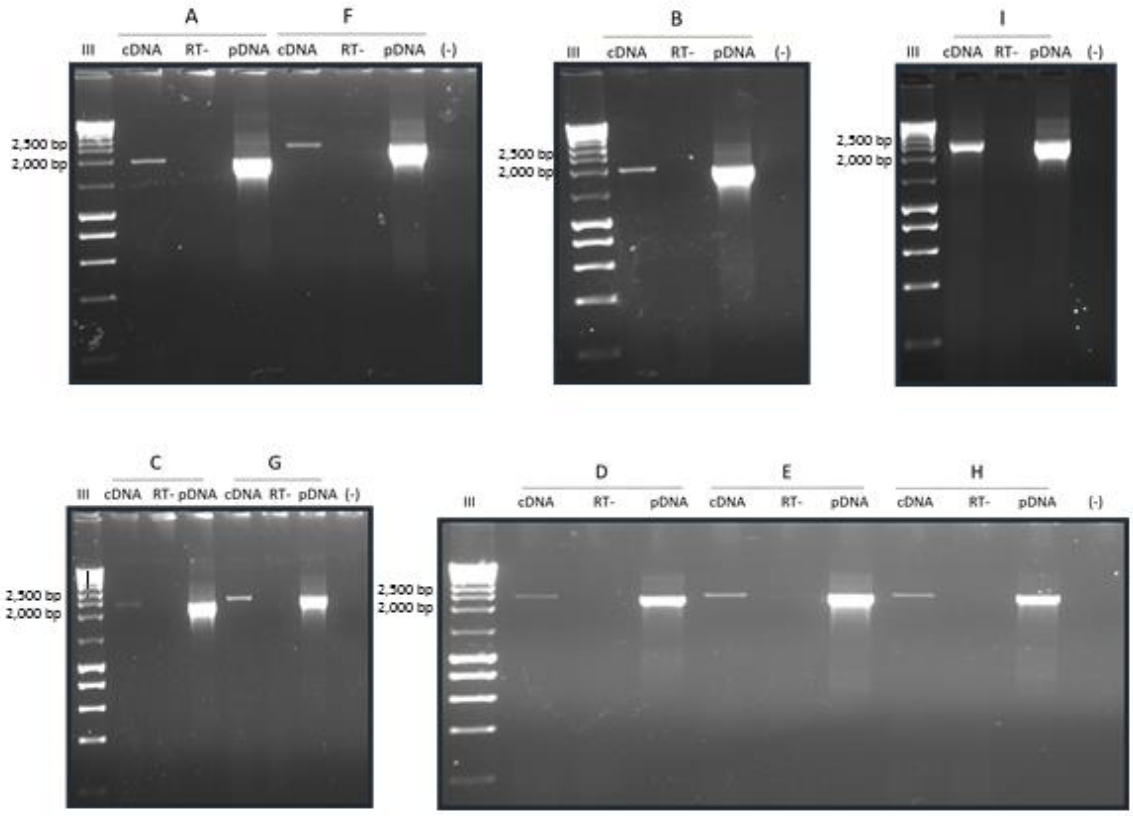


Figure 4. 13 — Analysis of the bicistronic mRNAs by RT–PCR using primers' set II.B reveals the correct full-length DNA fragment, confirming the full-length expressed mRNA. A, B, C, D, E, F, G, H and I are, respectively, the RT–PCR of *PR_F*, *PR_HBB_F*, *PR_MYC_F*, *PR_d160_F*, *PR_5'd160_F*, *PR_5'd160R175H_F*, *PR_5'd160R248Q_F*, *PR_5'd160R273H_F* and *PR_5'd160R282W_*F bicistronic mRNAs. III corresponds to NZYLadderIII DNA molecular weight ladder (NZYTech); RT- indicates the PCR amplification reaction without cDNA synthesis step, proving no DNA contamination occurred in the cDNA sample. The (-) lanes contain the PCR negative controls. Fragments size: 1,848 base pairs (bp) (A), 1,868 bp (B), 2,198 (C), 2,317 bp (D) and 2,397 bp (E, F, G, H and I). In this case, fragments amplified from pDNA do not include the chimeric intron, thus pDNA fragments have the same size as the corresponding cDNA. RT–PCR conditions: 3 µL of cDNA synthesis reaction; 1 mM of Mg^{2+} at final PCR reaction; PCR program: 95 ˚C for 3 min (initial denaturation), [95 ˚C for 30 s (denaturation), 57 ˚C for 30s (annealing), 72 ˚C for 2 min (extension)] x 35, 72 ˚C for 10 min (final extension).

The performed RT–PCR analysis shows just one fragment with the full-length size, which seems to indicate that only one transcript is produce from the bicistronic constructs. Thus, it seems that the obtained FLuc/RLuc relative expression bigger than 1 for *PR_d160_F* and *PR_5'd160R175H_F* bicistronic constructs, when compared to that from *PR_5'd160_F*, is not a consequence of alterative splicing and may be due to IRES-dependent translation. However, mRNA integrity analysis by RT–PCR is not sufficient to rule out alternative splicing events, since we can only estimate fragments size in the agarose gel. Accordingly, alternative splicing with removal of small DNA portions cannot be excluded. Thus, we should have extracted the obtained PCR bands to confirm the whole sequence of both amplified fragments by Sanger sequencing. Besides, designing siRNAs for RLuc and FLuc mRNAs

and performing knockdown of RLuc and Fluc individually can also complement our results. Mutual reduction of RLuc and FLuc expression, when separately silencing RLuc and FLuc expression, would also suggest that both luciferases are generated from the same transcript, excluding alternative splicing events.

At the end, in the presence of alternative splicing, *in vitro* transcription of the bicistronic mRNAs, followed by mRNA transfection, could be done to abolish this interfering event. However, some issues can also appear with this strategy. Evidence in the literature states that some IRES-containing mRNAs are only active if the mRNA is generated in the nucleus via RNA Pol II.[103] This suggests that a "nuclear experience" is required to activate some IRES.[103] A possible explanation is that some nuclear RNA-binding proteins may act as ITAFs, and therefore some protein–mRNA interactions must occur in the nucleus, followed by export to the cytoplasm and IRES-mediated translation.[103] This is observed for *XIAP* mRNA, since two known *XIAP* ITAFs that positively modulate its expression through the IRES element are the heterogeneous nuclear ribonucleoproteins C1 and C2 (hnRNPC1 and hnRNPC2).[104] Other cellular IRES such as *c-myc* IRES, the positive control used in this thesis, also require a nuclear event to ensure efficient non-canonical translation initiation.[105] This need is not fully understood. Nevertheless, as cellular IRES are less structured than viral IRES and may switch between cap-dependent and cap-independent translation, depending on cellular conditions, it is hypothesized that nuclear ITAFs could have an impact on the conformational state of the IRES element, thus assisting the switch between canonical and non-canonical mechanisms of translation initiation.[53] Indeed, stable RNA structures may difficult cap-dependent translation while favoring IRES-dependent mechanisms.[53] Consequently, if *in vitro* transcription is required for IRES activity assessment, a system in the presence and absence of a nuclear cellular extract should be generated, to evaluate the need for the "nuclear experience" and avoid false-negative results.

4.1.3.2 FLuc expression from all tested bicistronic constructs does not seem to be a consequence of cryptic promoter activity

The presence of a cryptic promoter can also lead to the transcription of aberrant bicistronic and/or monocistronic mRNAs, which, in turn, can express, cap-dependently, for an enzymatically active FLuc. Thus, this event can contribute to the appearance of false positives on IRES element identification. In fact, there are many reports in which a putative IRES was described, but, after cryptic activity assessment, cloned sequences proved to function as a promoter.[106] It should be noted that the presence of a cryptic promotor in the bicistronic construct does not exclude the capacity of the sequence cloned upstream FLuc mediate IRES-mediated translation; however, it requires a different experimental approach to validate an IRES in such conditions; for instance, the *in vitro* transcription of capped and polyadenylated bicistronic mRNAs.[83] In this situation, the aforementioned limitations of *in vitro* transcription should be taken in consideration.

To rule out the presence of cryptic promoters in the bicistronic constructs used in this thesis, we compared RLuc and FLuc expression from *PR_d160_F* and *PR_5'd160_F* with that from the respective promoterless bicistronic constructs—*P-R_d160_F* and *P-R_5'd160_F*. Promoter-containing *PR_F* and *PR_MLH1_F* [xiv] and the correspondent promoterless counterparts—*P-R_F* and *P-R_MLH1_F* [xv]—were, respectively, used as the negative and positive controls for cryptic promoter activity (Figure 4.14).

[xiv] Bicistronic constructs containing human *MLH1* 5' UTR, which has been described to include a cryptic promoter.[107,108]

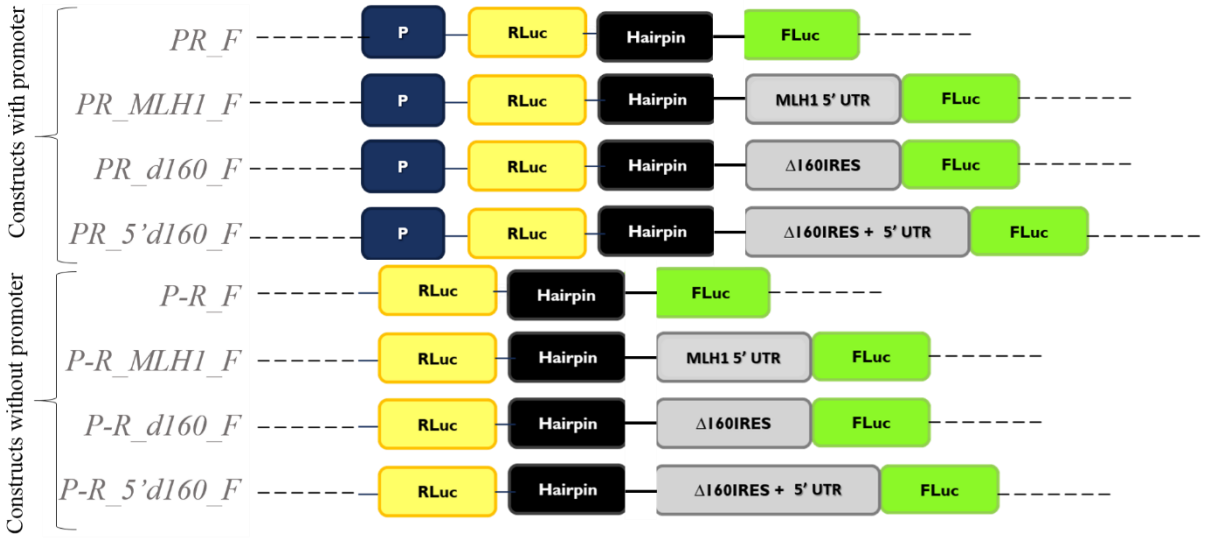


Figure 4. 14 — Bicistronic constructs used to check the presence of a cryptic promoter in Δ160p53 coding sequence or in its 5' UTR. RLuc is Renilla luciferase cap-dependent translated cistron (yellow box) and FLuc is firefly luciferase cap-independent translated cistron (green box). The blue box with the P letter represents the SV40 promoter. Grey boxes represent the cloned sequences upstream FLuc start codon. *PR_F* is the empty vector and *PR_MLH1_F*, the *MLH1* 5' UTR-containing vector, is the positive control for the presence of cryptic promoters. *PR_d160_F* corresponds to Δ*160p53* IRES-containing vector, which contains the first 432 nucleotides of Δ160p53 coding sequence. *PR_5'd160_F* vector contains Δ*160p53* IRES (first 432 nucleotides of Δ160p53 coding sequence) and its 5' UTR (all Δ133p53 coding sequence up to Δ160p53 start codon, excluding Δ133p53 start codon). *P-R_F*, *P-R_MLH1_F*, *P-R_d160_F* and *P-R_5'd160_F* are the counterpart promoterless bicistronic constructs.

HeLa cells were co-transfected with 1.5 µg (total DNA) of each of the aforementioned constructs and the β-galactosidase-encoding plasmid (pSV-β-Galactosidase Control Vector by Promega, a control vector for monitoring transfection efficiencies of mammalian cells). Twenty-four hours later, cells were lysed and bioluminescence assays were performed. Relative RLuc and FLuc expression values were obtained by normalizing each of them with those from β-galactosidase-encoding plasmid. Regarding relative RLuc/β-galactosidase expression, we expected higher levels in constructs containing SV40 promoter than in those it has been removed. Accordingly, as we can see in Figure 4.15, RLuc/β-galactosidase ratios from promoterless constructs are virtually inexistent. As far as relative FLuc expression is concerned, we observe a significant increase in its levels from *PR_MLH1_F*, compared to those from *PR_F*, as expected, due to the presence of a cryptic promoter in *MLH1*.[107,108] In fact, relative FLuc/β-galactosidase expression levels from promoterless *MLH1* 5' UTR-containing plasmid are significantly greater than those from *P-R_F*, the empty promoterless construct. For *P-R_d160_F* and *P-R_5'd160_F* promoterless constructs, FLuc/β-galactosidase ratio is not statistically different from that of *P-R_F*, which may suggest that Δ160p53 coding sequence and its 5' UTR do not contain any sequence capable of promoting transcription, with consequent formation of monocistronic mRNAs. Nevertheless, standard deviations of relative FLuc/β-galactosidase expression from *P-R_d160_F* and *P-R_5'd160_F* indicate a large dispersion of relative FLuc/β-galactosidase expression from *P-R_d160_F* and *P-R_5'd160_F* at each individual experiment, when compared to the mean. Therefore, more experiments, followed by outlier's exclusion, are required to guarantee the inexistence of cryptic promoter activity. Moreover, as for alternative splicing evaluation, individual knockdown of RLuc and FLuc can also complement these results, since mutual reduction of RLuc and FLuc expression, when separately silencing RLuc and FLuc expression, would further support that both luciferases are generated from the same transcript, excluding cryptic promoter activity.

Concluding and considering all the obtained results in sections 4.1.3.1 and 4.1.3.2, cryptic promoter activity and alternative splicing are unlikely causes for FLuc expression from the bicistronic

constructs. Notwithstanding, some experiments are still required to prove that the increase in FLuc expression from *PR_d160_F* and *PR_5'd160R175H_F* bicistronic constructs is due to non-canonical translational initiation mechanisms.

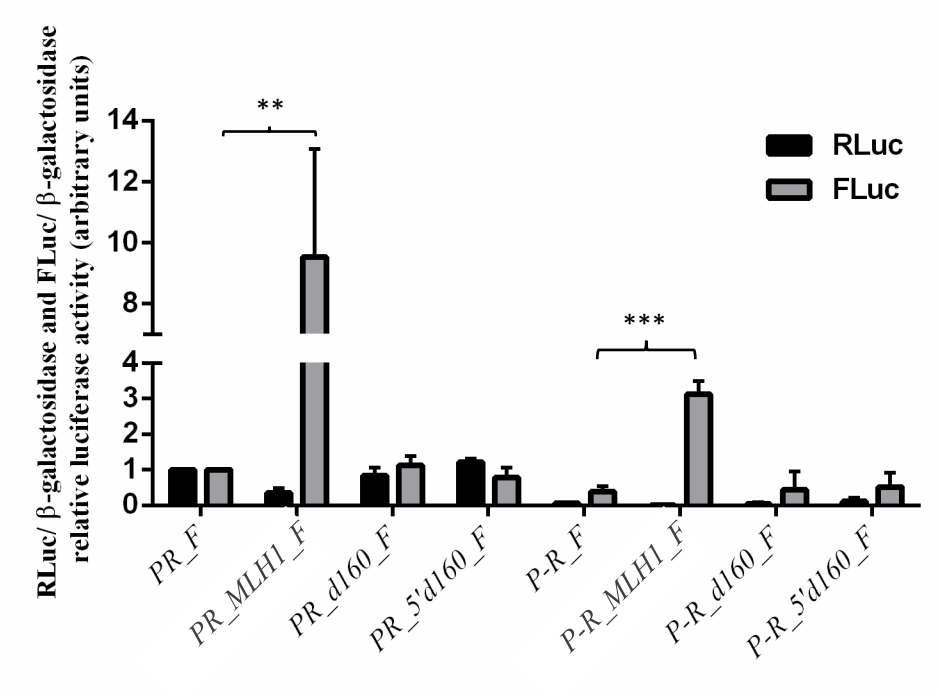


Figure 4. 15 — The bicistronic constructs *PR_d160_F* and *PR_5'd160_F* do not seem to have cryptic promoter activity. HeLa cells were transfected with promoter-containing constructs (*PR_F*, *PR_MLH1_F*, *PR_d160_F* and *PR_5'd160_F*) or promoterless constructs (*P-R_F*, *P-R_MLH1_F*, *P-R_d160_F* and *P-R_5'd160_F*) and co-transfected with β-galactosidase-encoding plasmid (pSV-β-Galactosidase Control Vector by Promega, a control vector for monitoring transfection efficiencies of mammalian cells). Relative RLuc and FLuc expression values (black and grey bars, respectively) were obtained by normalizing each of them with those from β-galactosidase-encoding plasmid, all measured by luminometry assays. Presented data are the result of at least three independent experiments. Asterisks (*) indicate statistical significance. **P<0.01, ***P<0.001

4.2 Unveiling new *p53* ITAFs: cloning of p53 isoforms in a 12 *MS2* repeat-containing vector

FL-p53 and Δ40p53 isoforms have been widely studied in terms of internal initiation mechanisms.[27,62,68] In fact, their IRES structures are already defined as well as some of the auxiliary proteins that mediate their non-canonical translation, such as Annexin A2, PSF and DAP5.[27,62,68] However, unlike longer p53 isoforms, the fact that Δ160p53 expression is mediated through an IRES element was not known until recently.[77] It was just in 2013 that a putative Δ*160p53* IRES located downstream AUG160 start codon was identified in our laboratory.[77] For that reason, the mechanisms of IRES-mediated translation of Δ160p53 are only now being elucidated and some questions still need an answer. For instance, the ITAFs that mediate Δ160p53 cap-independent translation are still not known. Therefore, in this thesis we also aimed to identify the proteins that may regulate Δ160p53 IRES-dependent translation, through direct or indirect interaction with the Δ*160p53* mRNA.

One way to identify auxiliary proteins that interact with a specific mRNA is using the MS2 system.[109] This system takes advantage of MS2 bacteriophage RNA and coat protein.[109] *MS2* RNA folds into a hairpin loop structure, due to base pairs complementarity, which is recognized by MS2 coat protein.[109] This interaction is very specific, and because of that, this system has been broadly used for the study of multiple steps of the eukaryotic mRNA life cycle, as it allows the identification of new RNA–protein complexes and new mRNA–miRNA interactions.[109–111] The experimental approach consists in tagging mRNAs with *MS2* hairpins and then affinity-purify *trans*-binding factors (RNA-binding proteins and non-coding RNAs) associated with the *MS2*-tagged mRNA.[112] Three experimental steps can be defined. The first involves the construction of two plasmid vectors (Figure 4.16), followed by their co-transfection.[111] The first plasmid expresses a chimeric RNA containing the mRNA of interest in which several *MS2* RNA hairpins motifs (usually 12 or 24 tandem *MS2* RNA hairpin loops) were introduced.[111] As MS2 coat protein binds cooperatively to adjacent hairpins, the use of multiple *MS2* RNA hairpins is desired to favor *MS2* hairpin–MS2 coat protein interactions.[113] The second plasmid expresses the MS2 coat protein, usually fused to a tag.[111]

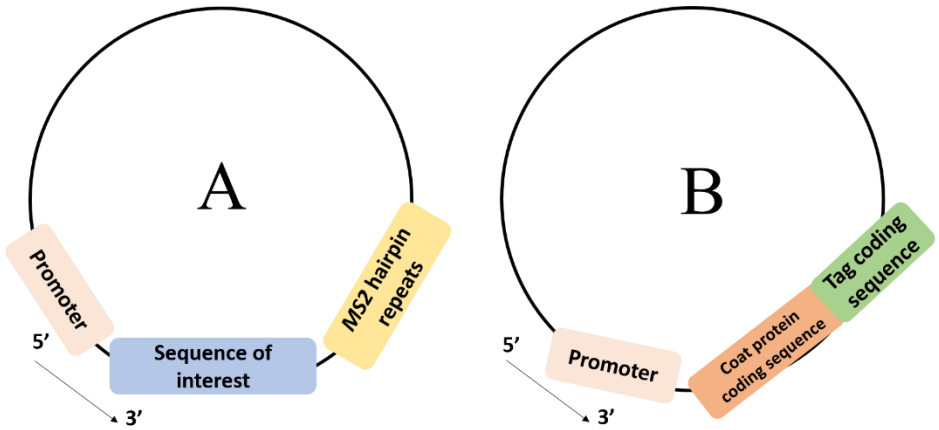


Figure 4.16 — The two plasmid vectors necessary for the MS2 system. "A" indicates the plasmid that expresses a chimeric RNA containing the mRNA of interest in which several *MS2* RNA hairpins motifs (usually 12 or 24 tandem *MS2* RNA hairpin loops) were introduced. "B" corresponds to the plasmid containing the coat protein coding sequence, which is usually fused to a tag. Pink, blue, yellow, orange and green boxes represent, respectively, the promoter, the sequence of interest, the several *MS2* RNA repeats, the coat protein coding sequence and the tag coding sequence.

After co-transfection with these two vectors, followed by cell lysis 24–48 h later, an immunoprecipitation assay can be performed, using an MS2 coat protein monoclonal antibody immobilized in a solid support (for instance, agarose recombinant protein G-containing beads).[110] Coat protein immunoprecipitation using a monoclonal antibody that binds to the tag can also be done.[110] Co-

immunoprecipitation of the chimeric RNA containing the mRNA of interest fused with several *MS2* RNA hairpins is expected, due to *MS2* hairpin–MS2 coat protein interaction.[109–112] Then, if the aim is to detect other RNAs that interact with the mRNA of interest, a targeted screen by reverse transcription and real-time quantitative PCR can be done.[109–112] On the other hand, if the aim is the characterization of ribonucleoprotein complexes formed on a given mRNA of interest, as in our purpose, mass spectrometry can be performed.[110]

Considering the aforementioned information, the first step to identify new ITAFs that regulate Δ*160p53* IRES using the MS2 system, is the generation of a construct with p53's sequence of interest upstream of *MS2* RNA repeats. The empty vector with twelve *MS2* hairpin repeats (*p_12MS2*) was already available in the laboratory.[xv] Then, we decided to clone four distinct p53 sequences upstream *MS2* RNA repeats: full-length p53 coding sequence, Δ133p53 coding sequence with part of its UTR[xvi], Δ160p53 coding sequence with its 5' UTR[xvii] and *Δ160p53* IRES[xviii] with its 5' UTR, generating, respectively, *pFLp53_12MS2*, *p5'd133p53_12MS2*, *p5'd160p53_12MS2* and *p5'd160IRES_12MS2* monocistronic constructs. As referred before, *TP53* gene allows the transcription of full-length mRNAs that are responsible for FL-p53, Δ40p53 and Δ160p53 expression, and allows the transcription of a shorter mRNA that expresses Δ133p53 and Δ160p53 as well.[68,69] Therefore, since we aim the identification of proteins that may regulate Δ160p53 non-canonical translation, our central constructs are *p5'd133p53_12MS2* and *pFLp53_12MS2*, as they are the ones that most resembles the naturally occurring p53 transcripts. Besides Δ*160p53* IRES, the *pFLp53_12MS2* construct also contains the Δ*40p53* IRES. This IRES, besides regulating Δ40p53 expression, can also activate FL-p53 expression during endoplasmic reticulum stress conditions, which are known to impair cap-dependent translation through eIF2α phosphorylation.[68,114] Thus, *pFLp53_12MS2* construct will allow us to compare proteins that regulate IRES-mediated translation of longer p53 isoforms (FL-p53 and Δ40p53) with proteins that regulate cap-independent translation of Δ160p53 shorter p53 isoform. The *p5'd160p53_12MS2* and *p5'd160IRES_12MS2* constructs along with *p5'd133p53_12MS2* will permit to localize the RNA regions in which the identified proteins directly interact.

Four cloning strategies, that will be briefly described below, were performed but none provided successful results. The first one focused on *EcoRI* and *BglII* restriction sites for p53 sequences insertion between CMV promoter and MS2 repeats (Figure 4.17).

[xv] Courtesy of Doctor Carmo Fonseca, Instituto de Medicina Molecular, Lisbon, Portugal.
[xvi] Last 21 nucleotides of the 5' UTR.
[xvii] Δ133p53 coding sequence up to AUG160, excluding Δ133p53 start codon.
[xviii] Δ160p53 coding sequence up to nucleotide 432.

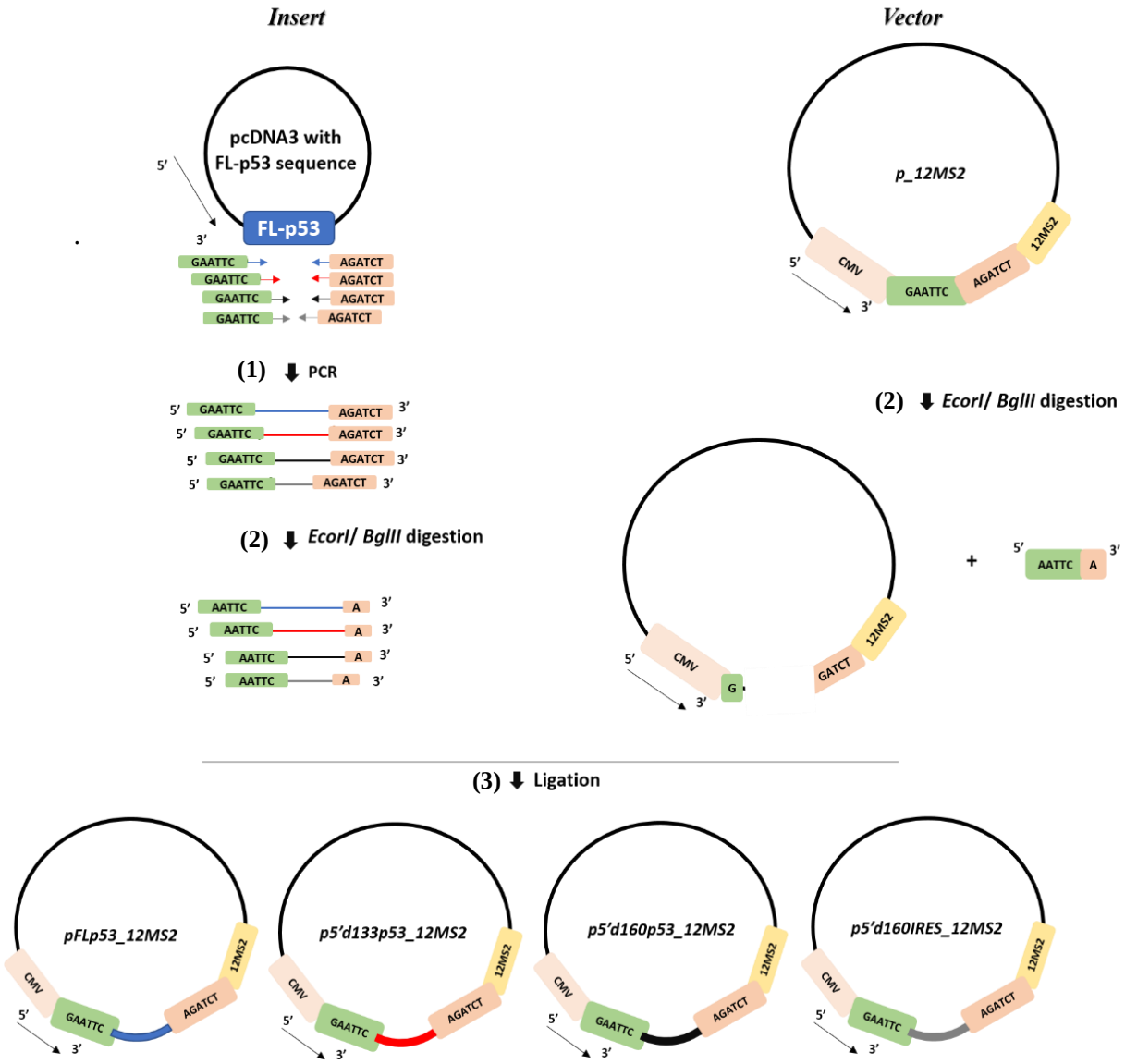


Figure 4. 17— The first cloning strategy was based on *EcoRI* and *BglII* restriction sites. **(1)** p53 sequences were amplified using primers containing linkers for *EcoRI* (forward primers) and *BglII* (reverse primers). **(2)** Afterwards, both PCR products and *p_12MS2* empty vector were digested with *EcoRI* and *BglII*. **(3)** Then, digested PCR products and empty vector were ligated using a vector/insert molar ratio of 1:3. *E. coli* competent transformation was performed and Bacteria were cultivated in LB agar plates supplemented with ampicillin. FL-p53 corresponds to full-length p53 sequence on pCDNA3 vector; ——— , ——— , ——— and ——— are full-length p53 coding sequence, Δ133p53 coding sequence with part of its 5' UTR , Δ160p53 coding sequence with its 5' UTR and *Δ160p53* IRES with its 5' UTR, respectively; →← and →← , →← and →← , →← and →← , →← and ← are, respectively, the forward and reverse primers used to amplify full-length p53 coding sequence, Δ133p53 coding sequence with part of its 5' UTR , Δ160p53 coding sequence with its 5' UTR and *Δ160p53* IRES with its 5' UTR; CMV is the promoter of *p_12MS2*; GAATTC and AGATCT are the restriction sites for *EcoRI* and *BglII*, respectively; G and AATTC correspond to *EcoRI* restriction sites after digestion; A and GATCT are the restriction sites of *BglII* after digestion; 12MS2 corresponds to the twelve repeats of *MS2* sequence.

After transforming competent bacteria with the ligation products, random colonies were selected for a colony screening PCR targeting the inserted sequence (Figure 4.18). To detect *pFLp53_12MS2*, amplification of a 1,201 base pairs (bp) fragment was expected. Two of the four tested colonies (C and D; Figure 4.18) seem to contain this construct, since a fragment between 1,000 and 1,400 bp is detected. Regarding *p5'd133_12MS2* vector, we expected an 826 bp fragment amplification in positive colonies. As the blank from PCR shows some DNA contamination between 800 and 1,000 bp, we only consider

as positives the colonies in which a more intense band is observed (A and D; Figure 4.18). For *p5'd160p53_12MS2* detection, we expected an 802 bp fragment amplification, whereas for *p5'd160IRES_12MS2*, a 529 bp fragment was the desired. For *p5'd160p53_12MS2* detection, although PCR blank control contamination between 800 and 1,000 bp, all colonies seem to be positive, since they present a more intense band than that observed for PCR blank control. On the other hand, in *p5'd160IRES_12MS2* detection, only colony C seems to be positive, displaying a fragment between 600 and 800 bp (Figure 4.18).

Notwithstanding, we should take in consideration that *p_12MS2* empty vector (lane P; Figure 4.18), used on colony screening PCR as a negative control, shows unspecific amplifications. Moreover, for *p5'd133_12MS2* and *p5'd160p53_12MS2* detection, a fragment with a similar size to that expected for the presence of *p5'd133_12MS2* and *p5'd160p53_12MS2* is observed. Unspecific amplifications can be explained by unspecific primer binding, as a consequence of low annealing temperature or excess of Mg^{2+} in the PCR reaction, for instance. However, the presence of a fragment with a similar size to that expected for the presence of the aforementioned vectors indicates primer contamination with a p53 sequence-containing vector. As the primer aliquots used for colony screening PCR were the same used to amplify p53 sequences of interest (Step 1 from Figure 4.17), primer's contamination with the vector used as template for p53 sequences of interest amplification, seems to be a reasonable explanation. In fact, this is corroborated by PCR blank control contamination in the colony screening PCR for *p5'd133_12MS2* and *p5'd160p53_12MS2* detection.

Although not sure that the positive colonies are in fact positive, we cultured them in liquid medium. Then, plasmid DNA was extracted and sent for sequencing (data not shown), which, in turn, revealed the presence of *p_12MS2* empty vector. Picking more colonies for a new colony screening PCR or repeating this first cloning strategy from the beginning continued to produce negative results.

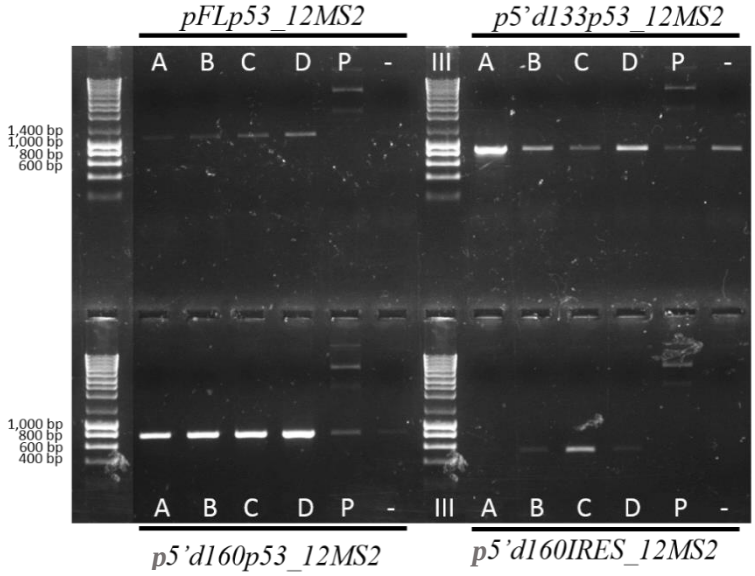


Figure 4. 18 — Colony screening PCR of constructs containing FL-p53 coding sequence, Δ133p53 coding sequence with part of its 5' UTR, Δ160p53 coding sequence with its 5' UTR and *Δ160p53* IRES with its 5' UTR shows positive colonies. *pFLp53_12MS2*, *p5'd133p53_12MS2*, *p5'd160p53_12MS2* and *p5'd160IRES_12MS2* monocistronic constructs contain, respectively, FL-p53 coding sequence, Δ133p53 coding sequence with part of its 5' UTR, Δ160p53 coding sequence with its 5' UTR and *Δ160p53* IRES with its 5' UTR, fused with 12 *MS2* RNA hairpins. A, B, C and D lanes correspond to the tested colonies and the P lanes correspond to *p_12MS2* empty vector negative control. A blank control without DNA (-) was prepared for each pair of primers used for each PCR amplification. III represents NZYLadderIII DNA molecular weight ladder (NZYTech). bp, base pair.

One possible explanation for these negative results is the fact that *EcoRI* and *BglII* restriction sites are next to each other in *p_12MS2* empty vector. Although there is no overlapping, no nucleotide

separates these two restriction sites. Therefore, when one of the two enzymes cleaves the DNA, a sticky end is produced and only one base pair separates the other enzyme restriction site from the unpaired end. It is known that restriction endonucleases have a low efficient cleavage when its restriction site is almost at the end of the DNA.[115] For *EcoRI*, 5 bp from the end are needed to achieve an efficient cleavage.[115] For *BglII*, 2 bp are enough.[115] Nevertheless, after digesting the empty vector with one of these two enzymes, the other do not have the minimum of base pairs required at the end for efficient cleavage. Thus, after *BglII* and *EcoRI* digestion, we might be in the presence of linearized *p_12MS2* that circularized when we performed the ligation reaction, due to 5' and 3' unpaired ends complementarity, leading to the obtained negative results.

Having this in mind, we moved to another cloning strategy, in which only one *BglII* digestion and an alkaline phosphatase treatment of the linearized vector were used to clone the p53 sequences of interest into the *p_12MS2* empty vector (Figure 4.19).

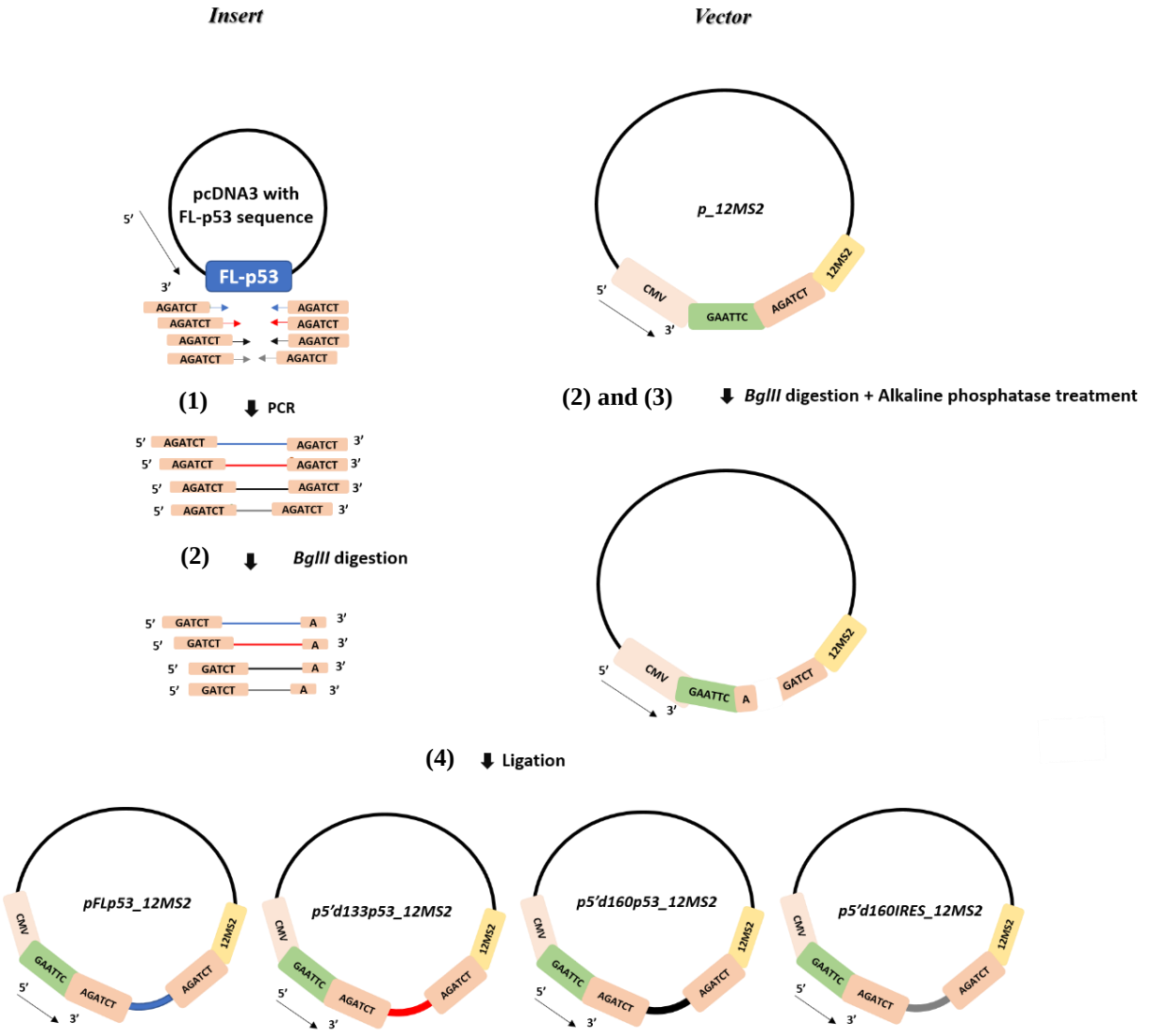


Figure 4. 19—The second cloning strategy was based on the *BglII* restriction site. **(1)** p53 sequences were amplified using primers containing linkers for *BglII* (forward and reverse primers). **(2)** Afterwards, both PCR products and *p_12MS2* empty vector were digested with *BglII*. **(3)** To avoid plasmid re-ligation, the phosphate groups on vector 5' -end were removed using an alkaline phosphatase. **(4)** Then, vector and inserts were ligated using a vector/insert molar ratio of 1:3. *E. coli* competent transformation was performed and Bacteria were cultivated in LB agar plates supplemented with ampicillin. FL-p53 corresponds to full-length p53 sequence on pCDNA3 vector; ——, ——, ——, and —— are full-length p53 coding sequence, Δ133p53 coding sequence with part of its 5' UTR , Δ160p53 coding sequence with its 5' UTR and *Δ160p53* IRES with its 5' UTR, respectively; → and ←, → and ←, → and ←

, → and ← are, respectively, the forward and reverse primers used to amplify full-length p53 coding sequence, Δ133p53 coding sequence with part of its 5' UTR , Δ160p53 coding sequence with its 5' UTR and and *Δ160p53* IRES with its 5' UTR; CMV is the promoter of *p_12MS2*; GAATTC and AGATCT are the restriction sites for *EcoRI* and *BglII*, respectively; A and GATCT are the restriction sites of *BglII* after digestion; 12MS2 corresponds to the twelve repeats of *MS2* sequence.

After transforming competent bacteria with the ligation reactions, we performed a colony screening PCR (Figure 4.20). The expected fragments for the presence of *pFLp53_12MS2, p5'd133_12MS2, p5'd160p53_12MS2* and *p5'd160IRES_12MS2* were the same as before. As we can see in Figure 4.20 for *pFLp53_12MS2* detection, all tested colonies seem to be positive, since a fragment between 1,000 and 1,400 bp is observed in A, B, C and D lanes. Regarding *p5'd133_12MS2* construct, colony A seems to be positive as it shows amplification of a fragment between 800 and 1,000 bp. Although B, C and D show an 800–1,000 bp fragment amplification, we can also see a shorter fragment amplification, which indicates that they may contain more than just the sequence of interest. In *5'd160p53_12MS2* detection, colonies A, B and C seem to be positive (800–1,000 bp fragment amplification). As colony D also has a lower fragment amplification, it was excluded. At last, in *p5'd160IRES_12MS2* detection, colonies A and C seem to be positive, with a 500–600 bp fragment amplification, and colonies B and D were excluded since they also show a higher fragment amplification.

However, as before, unspecific amplification of our *p_12MS2* empty vector is observed as well as PCR blank control contamination in *5'd160p53_12MS2* detection. Nevertheless, the putative positive colonies were cultured in liquid medium and, afterwards, plasmid DNA was extracted and sequenced. The sequencing results (data not shown) identify plasmid DNAs as the *p_12MS2*.

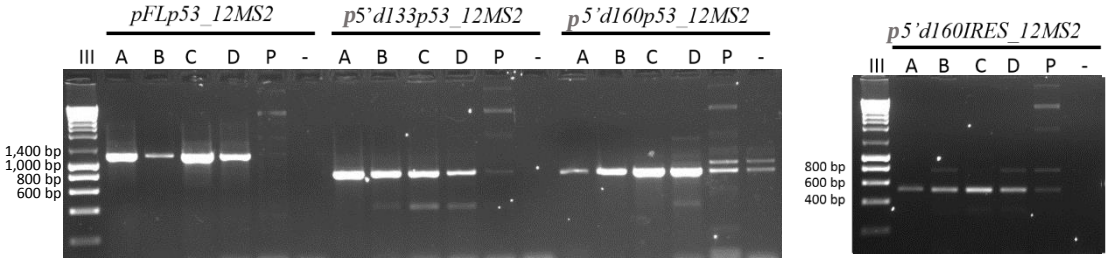


Figure 4. 20 — Colony screening PCR of constructs containing FL-p53 coding sequence, Δ133p53 coding sequence with part of its 5' UTR, Δ160p53 coding sequence with its 5' UTR and *Δ160p53* IRES with its 5' UTR shows positive colonies. *pFLp53_12MS2*, *p5'd133p53_12MS2*, *p5'd160p53_12MS2* and *p5'd160IRES_12MS2* monocistronic constructs contain, respectively, FL-p53 coding sequence, Δ133p53 coding sequence with part of its 5' UTR, Δ160p53 coding sequence with its 5' UTR and *Δ160p53* IRES with its 5' UTR, fused with 12 *MS2* RNA hairpins. A, B, C and D lanes correspond to the tested colonies and the P lanes correspond to *p_12MS2* empty vector negative control. A blank control without DNA (-) was prepared for each pair of primers used for each PCR amplification. III represents NZYLadderIII DNA molecular weight ladder (NZYTech). bp, base pair.

A crucial step in this cloning strategy is the removal of the vector's 5' –end phosphates, using an alkaline phosphatase, to prevent plasmid re-ligation without the insertion of p53 sequences of interest in the 12 *MS2* repeat-containing vector. This phenomenon can occur because we digested *p_12MS2* with only one enzyme (*BglII*) that produces sticky ends. Therefore, 5' and 3' -unpaired ends are complementary and can re-ligate. Thus, after obtaining negative cloning results, we checked the activity of the used alkaline phosphatase. First, we linearized *p_12MS2* vector with *BglII*. Then, we treated it with alkaline phosphatase and performed ligation reaction only with the vector. A non-treated vector was also ligated, to function as a control for alkaline phosphatase activity. *E. coli* DH5α competent bacteria transformation was done, followed by LB agar cultures in plates supplemented with ampicillin. Afterwards, we compared the number of colonies from both plates. While the plate with bacteria transformed with untreated vector had more than five hundred colonies, the plate with bacteria transformed with treated vector had less than one hundred colonies (data not shown). Thus, we can

conclude that alkaline phosphatase is working, however it is not very efficient, which might explain why we are obtaining *p_12MS2* empty vector. Therefore, we picked more colonies, seeking for a positive one among many containing *p_12MS2*. Yet, no positive results were achieved.

Consequently, we developed a third cloning strategy until we could purchase a new phosphatase. p53 sequences' amplification was accomplished following the same protocol as in strategy two (Figure 4.19). Then, while PCR products were only digested with *BglII*, *p_12MS2* vector was digested with *BsrGI* and *BglII*. Vector digestion produced two fragments: one with 4,540 base pairs (bp) and another with 649 bp. Afterwards, ligation between PCR products and vector longest fragment was accomplished. Then, a second ligation was performed to connect the ligation product of before with the vector shortest fragment. (Figure 4.21).

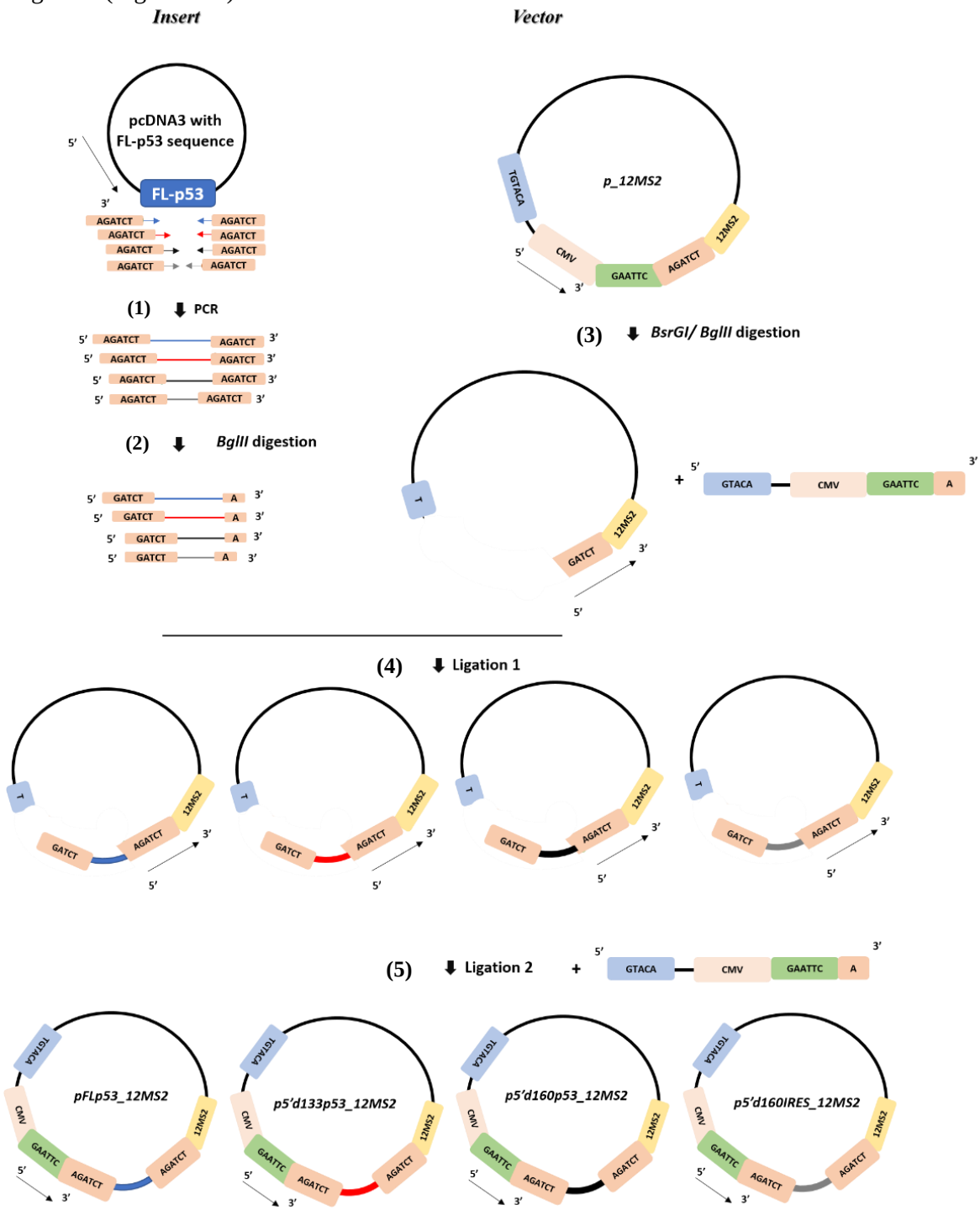


Figure 4. 21— The third cloning strategy was based on *BglII* and *BsrGI* restriction sites. **(1)** p53 sequences were amplified using primers containing linkers for *BglII* (forward and reverse primers). **(2)** Afterwards, PCR products were digested with *BglII*. **(3)** On the other hand, *p_12MS2* empty vector was digested with both *BsrGI* and *BglII*. Vector digestion produced two fragments: one with 4,540 base pairs (bp) and another with 649 bp. **(4)** p53 sequences and 4,540 bp fragment ligation, with a 1:3 vector/insert molar ratio was done. **(5)** Then, a second ligation was performed to connect the ligation product of before with the 649 bp fragment. A vector/insert molar ration of 1:3 was also used. E. coli competent transformation was performed and Bacteria were cultivated in LB agar plates supplemented with ampicillin. FL-p53 corresponds to full-length p53 sequence on pCDNA3 vector; ——— , ——— , ——— and ——— are full-length p53 coding sequence, Δ133p53 coding sequence with part of its 5' UTR , Δ160p53 coding sequence with its 5' UTR and *Δ160p53* IRES with its 5' UTR, respectively; → and ← , → and ← , → and ← , → and ← are, respectively, the forward and reverse primers used to amplify full-length p53 coding sequence, Δ133p53 coding sequence with part of its 5' UTR , Δ160p53 coding sequence with its 5' UTR and *Δ160p53* IRES with its 5' UTR; CMV is the promoter of *p_12MS2*; GAATTC , AGATCT and TGTACA are the restriction sites for *EcoRI*, *BglII* and *BsrGI*, respectively; A and GATCT are the restriction sites of *BglII* after digestion; T and GTACA are the restriction sites of *BsrGI* after digestion; 12MS2 corresponds to the twelve repeats of *MS2* sequence.

Random colonies of transformed bacteria were selected for a colony screening PCR (data not shown). We obtained some positive colonies that were put in liquid cultures for plasmid DNA extraction. Before sending the plasmid DNAs to sequencing, we performed a digestion screening with *BamHI* and *BsrGI* (Figure 4.22). While *BsrGI* restriction site is located upstream of CMV promoter, *BamHI* restriction site is located downstream of *MS2* hairpin motifs. No positive plasmid for *p5'd160p53_12MS2* was digested, since no positive colony grew on liquid culture, indicating that on colony screening PCR, those colonies were false positives. Regarding the digestion screening for the other constructs, if we were in the presence of *p_12MS2* empty vector, a 1,235 bp fragment. If we were in the presence of *pFLp53_12MS2*, *p5'd133p53_12MS2* or *p5'd160IRES_12MS2* constructs, a 2,423 bp, a 2,048 bp or a 1,751 bp fragment would be, respectively, observed. As we can see on Figure 4.21, all digested plasmid DNA correspond to *p_12MS2* empty vector, since a fragment shorter than 1,400 bp is detected (red arrows). Therefore, Sanger sequencing is not necessary as this screening allowed the exclusion of all obtained plasmids.

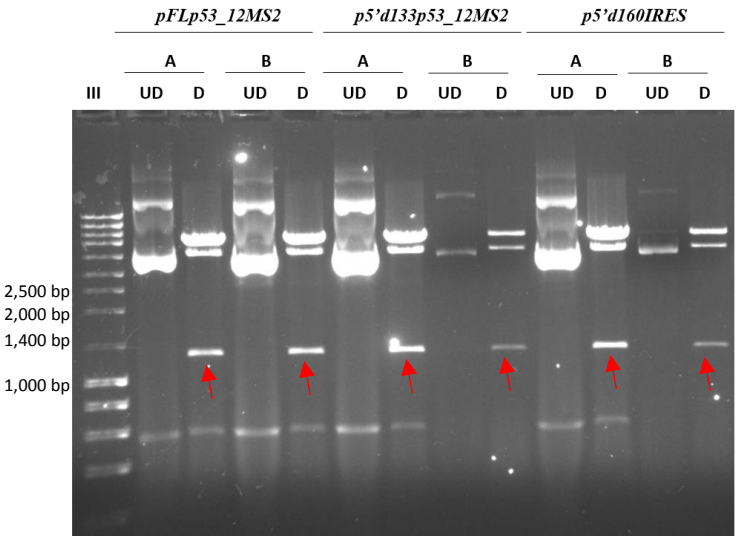


Figure 4. 22— Plasmid DNAs from positive colonies, digested with *BsrGI* and *BamHI*, correspond to *p_12MS2* empty vector. A and B are the putative positive plasmids obtained by DNA extraction of overnight liquid cultures of the colony screening PCR positive colonies. UD and D correspond to undigested and digested with *BsrGI* and *BamHI* lanes. *pFLp53_12MS2*, *p5'd160p53_12MS2* and *p5'd160IRES_12MS2* represent the monocistronic constructs that, respectively, contain FL-p53 coding sequence, Δ160p53 coding sequence with its 5' UTR and *Δ160p53* IRES with its 5' UTR, fused with 12 *MS2* RNA hairpins. III is NZYLadderIII DNA molecular weight ladder (NZYTech). Red arrows indicate the digested fragment used to distinguish between positive and negative plasmids for p53 sequences of interest. bp, base pair.

With another unsuccessful strategy, we changed the order of ligations. Therefore, p53 sequence of interest amplification, and digestion of PCR products and empty vector were done as before: while PCR products were only digested with *BglII*, *p_12MS2* vector was digested with *BsrGI* and *BglII*, which produced two fragments: one with 4,540 base pairs (bp) and another with 649 bp. However, in the fourth cloning strategy, we first ligate p53 sequences with the 649 bp vector fragment. Then, ligation between the obtained fragment with the 4,540 bp vector fragment was performed (Figure 4.23).

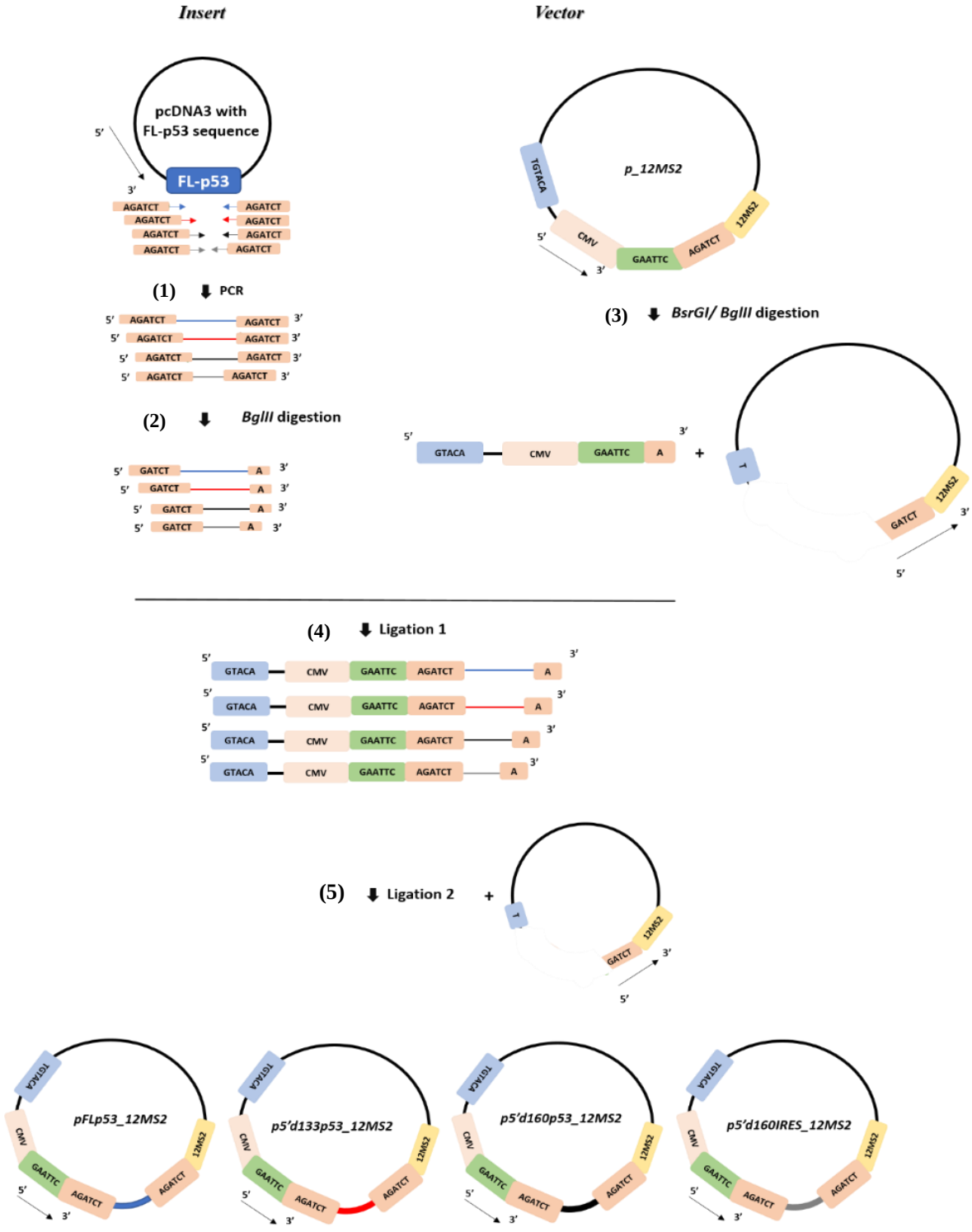


Figure 4. 23 — The fourth cloning strategy was based on *BglII* and *BsrGI* restriction sites. **(1)** p53 sequences were amplified using primers containing linkers for *BglII* (forward and reverse primers). **(2)** Afterwards, PCR products were digested with *BglII*. **(3)** On the other hand, *p_12MS2* empty vector was digested with both *BsrGI* and *BglII*. Vector digestion

with produced two fragments: one with 4,540 base pairs (bp) and another with 649 bp. **(4)** p53 sequences and 649 bp fragment ligation, with a 1:1 molar ratio was done. **(5)** Then, a second ligation was performed to connect the ligation product of before with the 4,540 bp fragment. A vector/insert molar ration of 1:3 was used. E. coli competent transformation was performed and Bacteria were cultivated in LB agar plates supplemented with ampicillin. FL-p53 corresponds to full-length p53 sequence on pCDNA3 vector; ——— , ——— , ——— and ——— are full-length p53 coding sequence, Δ133p53 coding sequence with part of its 5' UTR , Δ160p53 coding sequence with its 5' UTR and *Δ160p53* IRES with its 5' UTR, respectively; → and ← , → and ← , → and ← , → and ← are, respectively, the forward and reverse primers used to amplify full-length p53 coding sequence, Δ133p53 coding sequence with part of its 5' UTR , Δ160p53 coding sequence with its 5' UTR and *Δ160p53* IRES with its 5' UTR; CMV is the promoter of *p_12MS2*; GAATTC , AGATCT and TGTACA are the restriction sites for *EcoRI*, *BglII* and *BsrGI*, respectively; A and GATCT are the restriction sites of *BglII* after digestion; T and GTACA are the restriction sites of *BsrGI* after digestion; 12MS2 corresponds to the twelve repeats of *MS2* sequence.

As before, random colonies of transformed bacteria were selected for a colony screening PCR (data not shown). Positive colonies were cultured in liquid medium for plasmid DNA extraction. Before sequencing plasmid DNA, we performed a digestion screening with *BamHI* and *BsrGI*, as described previously. As we can see in Figure 4.24, for *p5'd133p53_12MS2* detection, plasmid E seems to be positive, since a fragment between 2,000 bp and 2,500 bp is observed (green arrow), which may correspond to the expected 2,048 bp fragment. The other plasmids present a fragment shorter than 1,400 bp (red arrows), which seems to correspond to the expected 1235 bp fragment of *p_12MS2* empty vector. Thus, only plasmid E was sequenced.

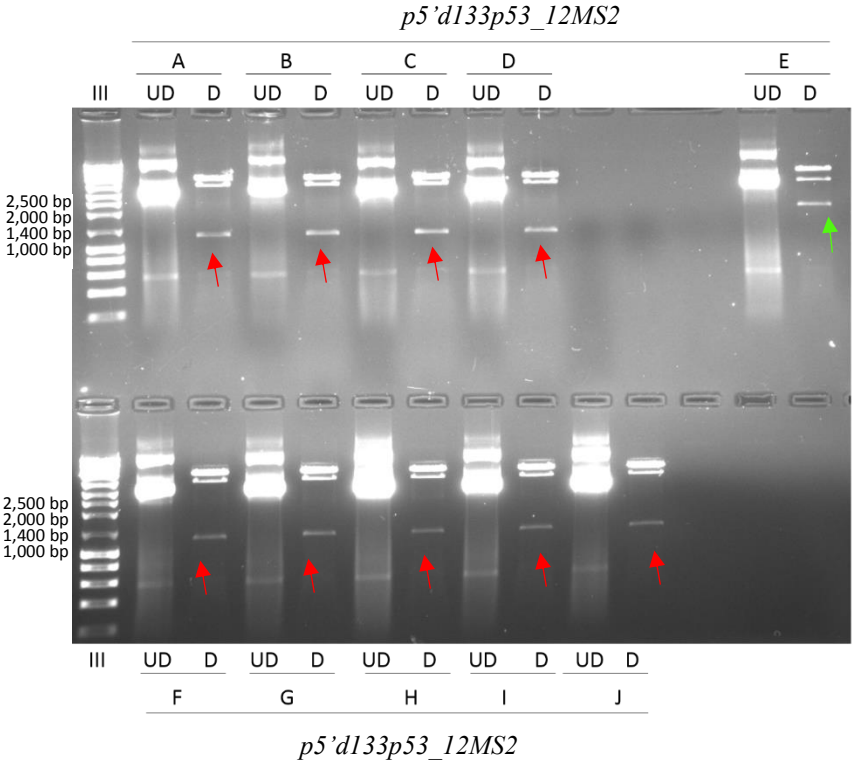


Figure 4. 24— Plasmid DNA E, digested with *BsrGI* and *BamHI*, seems to correspond to *p5'd133p53_12MS2*. A to J are the putative positive plasmids obtained by DNA extraction of overnight liquid cultures of the colony screening PCR positive colonies. UD and D correspond to undigested and digested plasmid with *BsrGI* and *BamHI* lanes. *p5'd133p53_12MS2* represents the monocistronic construct containing Δ133p53 coding sequence with part of its 5' UTR. III is NZYLadderIII DNA molecular weight ladder (NZYTech). Arrows indicate the digested fragment used to distinguish between positive (green) and negative (red) Δ133p53 coding sequence with part of its 5' UTR-containing plasmids. bp, base pair.

It is noteworthy that as we are only using *BglII* restriction site to clone p53 sequences into *p_12MS2*, we may obtain a construct in which p53 sequence is inserted with a 3' → 5' orientation and not with a 5' → 3' orientation, as we want. If fact, sequencing results of the positive plasmid showed

that p53 sequence is cloned backwards. For the other constructs, only *p_12MS2* empty vector was obtained. The presence of empty vector after transforming bacteria with the ligation products raises the possibility of inefficient digestion by restriction endonucleases and an inadequate ratio of insert to vector during the ligation. So, we also transformed bacteria with a ligation product in which only the vector was added. Therefore, this ligation reaction functioned as a control for the presence of empty vector. Indeed, colonies from the control ligation grew in LB agar plates and were one tenth of what we obtained in plates with transformed bacteria with our cloned products. Therefore, inefficient vector digestion may be part of the problem. We should also try to improve vector/insert molar ratios in the ligations. Another strategy can be to introduce 6 to 9 bp between *EcoRI* and *BglII* restriction sites by site-directed mutagenesis and then try the first cloning strategy again.

After obtaining the desired constructs, they will be co-transfected with the MS2 coat protein-containing plasmid, which will be followed by MS2 coat protein immunoprecipitation. Then, samples will be sent to mass spectrometry for possible new *p53* ITAFs identification. However, not all identified proteins can be considered *p53* ITAFs. Some of them might be interacting with tested *p53* mRNAs for other reasons, such as nuclear export. Therefore, silencing their expression individually and evaluate the effect in p53 isoform IRES translation will be needed to clarify their role in p53 IRES-mediated translation.

It would be of great relevance to identify common ITAFs of full-length and short p53 proteins IRES-mediated translation, as well as, specific ITAFs of each, through comparison of *pFL-p53_12MS2* and *p5'd133p53_12MS2* interacting proteins. Shared *p53* ITAFs may also contribute to IRES-mediated translation of other mRNAs. Therefore, further studies focused on them may lead not only to the identification of novel mRNAs with IRES-dependent translation but also to cancer-related mRNAs. Indeed, some ITAFs of *FL-p53* and Δ*40p53* IRES, such as PTB,[68] have a role in the regulation of IRES-mediated translation of other cellular mRNAs involved in cancer, as *cat-1* and *Bag-1*.[29,45 82,116,117] Moreover, hnRNP C1 and hnRNP C2 RNA-binding proteins regulate IRES-mediated translation of XIAP, c-myc and PITSLRE.[118,119] Once again, these mRNAs are frequently overexpressed in tumor cells, playing a crucial role in tumor progression.[120–122] Then, deregulation of ITAFs shared between cancer-related mRNAs could play a preeminent role in cancer development. Nevertheless, their part in tumor progression may not be linear. Indeed, PTB can also positively regulate IRES-mediated translation of mRNAs that have a reduced expression in cancer cells, as *Apaf-1*.[45,123] Besides, ITAFs can stimulate IRES activity of some mRNAs, while reducing IRES function of others.[45,48,124] Therefore, a huge lack of information still exists regarding the role of ITAFs in cancer and additional studies are required. We hope to unveil some knowledge on this matter after concluding our studies with the MS2 system. Besides identification of common and specific ITAFs of full-length and short p53 IRES-mediated translation, we also aim to understand the short *p53* transcript interactions with the identified ITAFs. In fact, *p5'd160p53_12MS2* and *p5'd160IRES_12MS2* constructs along with *p5'd133p53_12MS2* will narrow protein–mRNA interactions. ITAFs interacting with both *p5'd160p53_12MS2*- and *p5'd133p53_12MS2*-transcribed mRNAs, but not with *p5'd160IRES_12MS2* mRNA may indicate an interaction outside the IRES structure. On the other hand, identified *p5'd160IRES_12MS2* mRNA-interacting proteins could be the key factors in Δ160p53 IRES structure stabilization, since they directly interact with Δ*160p53* IRES and Δ*160p53* 5' UTR. Nonetheless, Δ*160p53* IRES activity should be the result of the combined action of both IRES-interacting proteins and Δ*133p53* 5' UTR-/Δ160p53 final coding region-interacting proteins. Besides, it is noteworthy that these ITAFs may not be just mRNA-binding proteins. Protein complexes can be formed for IRES secondary structure stabilization and ribosome recruitment, as occurs for IRES-mediated translation of *XIAP* mRNA.[27,29] Thus, distinction between proteins that directly bind to the mRNA and those that indirectly interact with *p53* shorter transcript will require extra studies. Additionally, as stated before, the mechanisms underlying ITAF function are still not known, but is believed they can remodel RNA

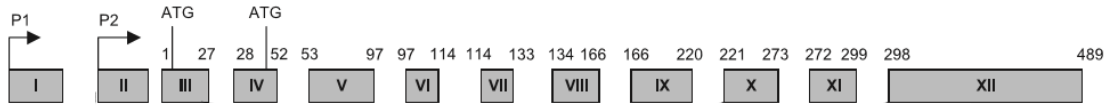
structures or work as a bridge between the mRNA and the ribosome.[27,29] Thus, clarification on *p53* ITAF function would be of great significance, since many new *p53* ITAFs may be shared by other IRES-containing mRNAs.

At the end, identification of new ITAFs regulating *p53* mRNA IRES-mediated translation, along with the study of cell stress conditions that stimulate their activity, will uncover new understanding on how IRES elements are regulated and how this cap-independent translation initiation mechanism can play crucial roles in cancer.

4.3 Unveiling new mRNAs regulated by Hdm2: immunoprecipitation assay optimization

Hdm2, the human homolog of murine Mdm2, is an E3 ubiquitin ligase that is responsible for the ubiquitination and consequent degradation of N-terminal-containing p53 isoforms.[125] The *HDM2* gene, composed of 12 exons, is regulated by two promoters.[125] The first one is located upstream exon I and the second one upstream exon II.[125,126] The first two exons of *HDM2* gene do not encode for any protein, and the first AUG is in exon III.[125,126] Thus, both promotors allow the expression of a full-length Hdm2 protein, with 90 kDa (Figure 4.25.1).[126] Besides FL-Hdm2, many other isoforms can be expressed, not only in normal tissues, but also in tumor cells.[126] They can be generated through alterative splicing and internal initiation at codon 52 (Figure 4.25.1).[126] Although more than 70 isoforms have already been identified, little is known about their function.[126] Three of these shorter isoforms—Hdm2-A, Hdm2-B and Hdm2-C—are commonly detected in various types of cancer.[126] Moreover, while FL-Hdm2 contains five functional domains, these three isoforms lack at least one of them, including the p53 binding domain.[126] Nevertheless, they maintain the ability to bind to FL-Hdm2, which means they can regulate FL-Hdm2 activity (Figure 4.25.2).[126]

(1)



(2)

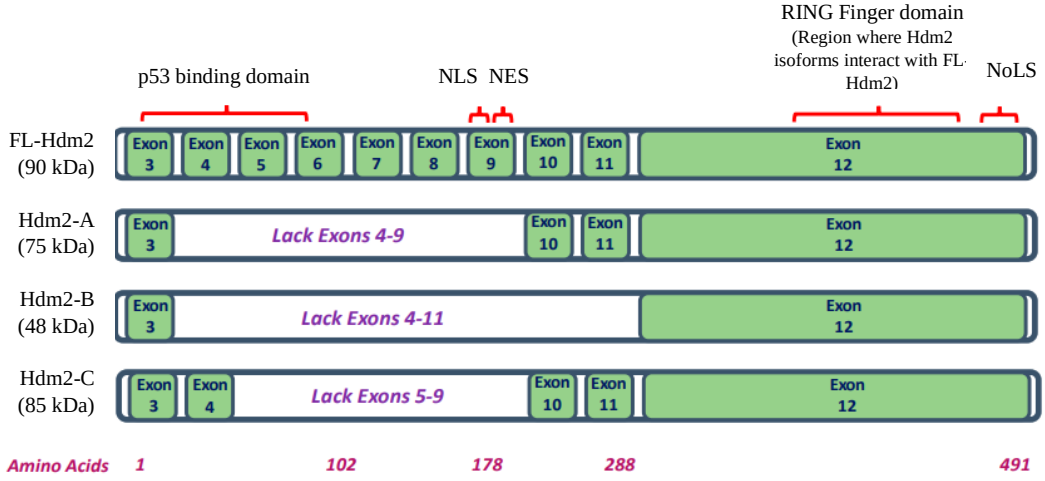


Figure 4. 25— *HDM2* gene, composed of 12 exons, can generate more than 70 Hdm2 isoforms. (1) *HDM2* gene is regulated by two promoters (P1 and P2), located upstream exons I and II, respectively. As the first two exons do not encode any protein, and the first AUG is in exon III, both promotors allow the expression of a full-length Hdm2 protein, with 90 kDa. The shorter Hdm2 isoforms are obtained through alterative splicing, aberrant splicing (between introns and exons) and internal initiation at codon 52. The roman letters represent the exons and the numbers above denote amino acid numbers. (2) FL-Hdm2 is composed of five domains. At its N-terminus is the p53 binding domain. The nuclear localization signal (NLS) and the nuclear export signal (NES) are within exon 9. At FL-Hdm2 C-terminus are the RING finger domain and the nucleolar localization signal (NoLS). Although Hdm2-A, Hdm2-B and Hdm2-C lack part of p53 binding domain, they maintain the RING finger domain, which allows the interaction of these Hdm2 shorter isoforms with FL-Hdm2. Adapted from Iwakuma *et al.*, 2003[125] and Saadatzadeh *et al.*, 2017[126].

Besides its role on p53 degradation, FL-Hdm2 can directly interact with full length *p53* mRNA. This interaction is mediated by the RING finger domain of Hdm2 and the 5' -terminus of full-length *p53* transcripts.[127] This interaction results in the induction of not only FL-p53 but also Δ40p53 protein.[128] Moreover, the RING domain of Hdm2 seems to interact with many other mRNAs, regulating their expression.[129] For instance, interaction between Hdm2 RING finger and *XIAP* mRNA leads to XIAP non-canonical translation through its IRES element.[129] Therefore, Hdm2 acts as an ITAF in XIAP IRES-dependent translation.[129] In addition, *VEGF*, *N-myc* and *Slug* mRNAs also interact with Hdm2.[130] Moreover, all these mRNAs have been associated with cancer progression.[27,53] Thus, since in malignant cells canonical translation is usually impaired,[27,53] mRNA translation regulation mediated by Hdm2 may be done through non-canonical mechanisms, including through IRES elements.

Having this in mind and knowing that understanding which key proteins are being translated under tumor microenvironment conditions, such as nutrient deprivation, hypoxia, genotoxic and oxidative stress, is an important step to understand the cellular mechanisms that regulate cancer cell survival and tumor progression, we defined the last aim of this thesis. Here, we aimed the identification of novel cancer-related mRNAs IRES-dependently translated during stress conditions, whose expression is regulated by Hdm2. In order to do that, we intended to perform co-immunoprecipitation (co-IP) of Hdm2 with its bound mRNAs, using Hdm2 antibodies, followed by RNA deep sequencing analysis.

The first step to achieve our goal is Hdm2 immunoprecipitation (IP) assay optimization, which can be assessed by Western blot analysis. Previous optimizations for Hdm2 IP assay were already performed in our laboratory, using 4B2 and SMP14 Hdm2 antibodies.[84] 4B2 recognizes an epitope in p53 binding site, located within Hdm2 amino acid number 19 and amino acid number 50 (Figure 4.26).[131] Thus, it can bind to FL-Hdm2 and Hdm2-C, but not to Hdm2-A and Hdm2-B.[131] 4B2 can also bind to other less studied isoforms containing the corresponding epitope.[131] On the other hand, SMP14 epitope locates between amino acid 154 and amino acid 167, integrating part of the nuclear localization signaling domain (Figure 4.26).[131] Thus, SMP14 only recognizes FL-Hdm2 and not the other three most common isoforms.[131] Previous immunoprecipitations were able to isolate some Hdm2 isoforms but not FL-Hdm2 (which should happen since both 4B2 and SMP14 recognize this isoform).[84] Besides, during Western blot analysis, none Hdm2 isoforms were detected on total lysate samples.[84] These samples corresponded to total protein input obtained from cleared cell lysates after lysis with IP buffer A and pellet removal.[84] Therefore, further optimizations on Hdm2 IP assay are needed.

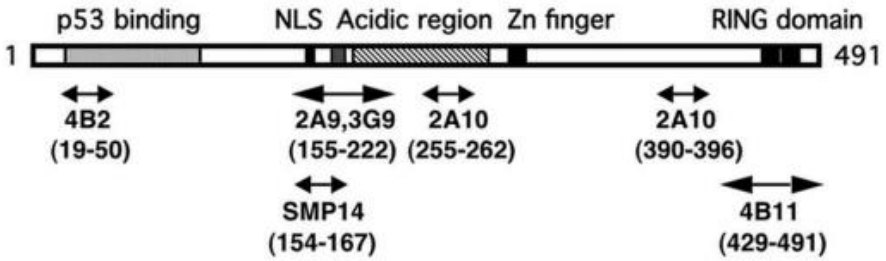


Figure 4. 26 — Diagram of Hdm2 and epitope locations for several Hdm2 monoclonal antibodies. The used antibodies 4B2 and SMP14 recognize 19–50 amino acids and 154–167 amino acids from Hdm2, respectively. Nuclear Localization Signal, NLS; Zinc finger domain, Zn finger. Adapted from Cheng, G. and Chen, J. (2011).[131]

Considering the aforementioned information, we started by transfecting 80 % confluent A549 cells with 4 µg of *pΔN6_HDM2*, which contains Hdm2 coding sequence. Then, 24 h later, 100 % confluent cells were harvested and lysed with IP buffer A, followed by Hdm2 immunoprecipitation, using either 4B2 or SMP14, both diluted 1:100, bound to agarose recombinant protein G-containing beads. A control sample without Hdm2 antibodies (MOCK) was also prepared to evaluate unspecific

binding to the beads. For Western blot analysis, we collected the total lysate (TL), containing total protein input obtained from cleared cell lysate; the flow-through (FT), with the proteins that did not bind to Hdm2 antibodies; and the immunoprecipitated sample (IP), with the proteins that were able to bind to Hdm2 antibodies. In order to detect Hdm2 on Western blot we used 4B2 monoclonal antibody (Figure 4.27). Therefore, on TL sample, we expected to detect all Hdm2 isoforms containing the epitope for 4B2, including FL-Hdm2 and Hdm2-C. On MOCK IP sample, we were not expecting to detect any protein, since this was a control for unspecific binding to the beads. In immunoprecipitated sample with SMP14, we were expecting to detect Hdm2 isoforms that simultaneously contain the SMP14 epitope and also the 4B2 epitope, since this was the antibody used on Western blot. Therefore, we expected to detect FL-Hdm2, but not Hdm2-A, Hdm2-B or Hdm2-C. In immunoprecipitated sample with 4B2, we expected to detect FL-Hdm2 and Hdm2-C. On FT samples for MOCK, SMP14 and 4B2, Hdm2 isoforms containing 4B2 epitope could appear in all these three lanes. However, we were expecting less band intensity on SMP14 and 4B2 lanes, since in these cases, Hdm2 isoforms should be essentially in IP samples. As we used mouse monoclonal antibodies in both immunoprecipitation and Western blot analysis, we were also expecting to detect the heavy and light chains of the antibodies used in IP, since during sample denaturation, antibodies detached from the beads. If heavy chains were detected, we would expected a 55 kDa band; as for light chains, we would expected a 25 kDa band.[132]

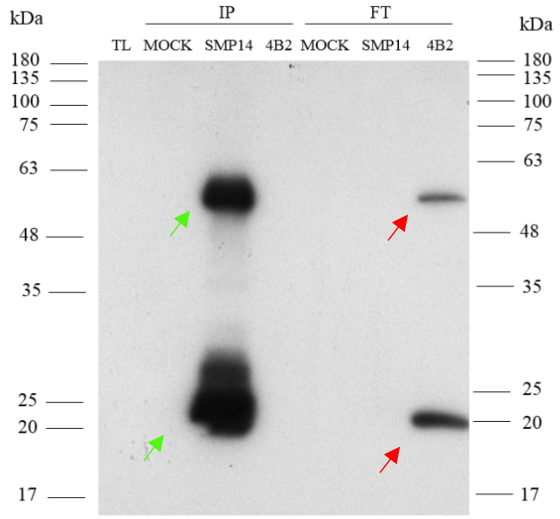


Figure 4. 27— SMP14 Hdm2 monoclonal antibody successfully bound to agarose recombinant protein G-containing beads, but no Hdm2 is detected by Western blot analysis. Immunoprecipitation of Hdm2 using either SMP14 or 4B2 antibodies diluted 1:100 from a cleared lysate of 100 % confluent A549 cells transfected with *pΔN6_HDM2* and lysed with IP buffer A. MOCK control was prepared through incubation of cleared lysate with beads, without any antibody incubation. For each immunoprecipitation, the total lysate of a 100-mm dish was used. Green arrows indicate SMP14 heavy and light chains; red arrows indicate 4B2 heavy and light chains. TL, total lysate; IP, immunoprecipitated samples; FT, flow through. NzyColour Protein Marker II was used. kDa, kilodalton.

As we can see in figure 4.27, SMP14 Hdm2 monoclonal antibody bound to the beads, since we detect its heavy and light chains in IP sample with that antibody (green arrows). On the other hand, 4B2 did not bind to the beads, as its chains only appear in FT sample (red arrows). A possible explanation for this can be the conformational changes in 4B2 antibody, due to freeze/thaw successive cycles, with a consequent loss of its binding capacity.[133] Furthermore, there was not any unspecific binding to agarose recombinant protein G-containing beads, since MOCK do not show any band in IP sample. However, we cannot detect any Hdm2 isoform, and because of that, we cannot assess immunoprecipitation efficiency. There are many reasons for no Hdm2 detection, such as low protein levels in our samples as

a result of inefficient transfection. Nevertheless, endogenous levels of Hdm2 should be detected. Indeed, Hdm2 is usually overexpressed in cancer lung cells like A549[134]. Therefore, lysis might have been inefficient, or our Western blot conditions still need to be optimized.

Then, we transfected 80 % confluent A549 cells with 1.5 µg of *pΔN6_HDM2* construct and tested lysis efficiency of IP buffer A and another IP lysis buffer (IP buffer B). Both IP buffers contain Nonidet-P 40 (Invitrogen) as the lysis agent. However, IP buffer B also contains glycerol. When we lyse a cell, we disrupt the cellular membrane, thus membrane hydrophobic regions become exposed.[135] These regions may destabilize proteins and promote their aggregation.[135] Glycerol helps in protein conformation stabilization, by inducing protein compaction and reducing protein flexibility.[135] In this experiment, we also used a non-transfected control and we also lysed cells with 5 x SDS sample buffer. This buffer contains SDS, which is considered a strong detergent with a great lysis capacity[136], thus samples lysed with this buffer were used to compare lysis efficiency of both tested IP buffers. We lysed cells 24 h post-transfection and evaluated its efficiency by Western blot. For Western blot optimization, we tested SMP14 Hdm2 monoclonal antibody. Moreover, we incubated the blotted PVDF membrane for 6 h with the anti-mouse secondary antibody, instead 1 h, as done before. We also tested a commercial ECL (Invitrogen), instead the homemade ECL used before. Commercial ECL allows us to expose our membranes up to 1 h, which contrasts to the 20-min maximum exposure of homemade ECL.

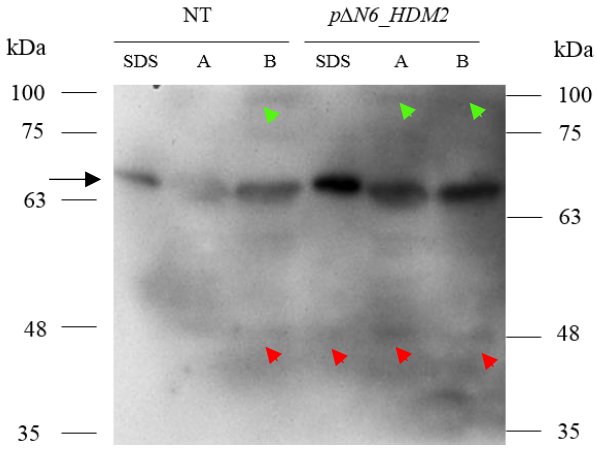


Figure 4. 28 — IP buffer B seems to be better than IP buffer A to lysate cells for Hdm2 detection. A549 cell lysis was performed using either 5 x SDS sample buffer (SDS lane), IP buffer A (A lane) or IP buffer B (B lane). Western blot analysis was performed using SMP14 Hdm2 monoclonal antibody diluted 1:500, overnight at 4 °C, followed by 6-h incubation with anti-mouse secondary antibody. Commercial ECL (Invitrogen) was used. Green arrows indicate 90 kDa FL-Hdm2; red arrows indicate 48 kDa Hdm2 isoform and the black arrow indicates a 63–75 kDa unidentified band. NzyColour Protein Marker II was used. NT, non-transfected; kDa, kilodalton.

Since we used SMP14 for Hdm2 detection in Western blot, we expected to detect FL-Hdm2, but not Hdm2-A, Hdm2-B and Hdm2-C. At Figure 4.28, we can see an intense band between 63 kDa and 75 kDa (black arrow). This may correspond to a less studied Hdm2 isoform or to cleavage products of FL-Hdm2 due to caspases activity. In fact, the resulting products migrate around 60 kDa.[137] Furthermore, FL-Hdm2 is hardly detected—a weak band appears only in non-transfected cells lysed with IP buffer B, and in transfected cells lysed with IP buffer A and B (green arrows). Therefore, the detection of cleavage products of FL-Hdm2 seems to be a good explanation for the most intense band (white box). We also detect a poorly expressed 48-kDa Hdm2 isoform in almost all lanes (red arrows). However, it cannot correspond to Hdm2-B, since this isoform does not include an epitope recognized by SMP14 antibody. To choose the best conditions for further experiments, we focused on the 63–75 kDa band (black arrow). Transfecting cells seems to be the better option for Hdm2 detection, since we

see an increase in this isoform levels in transfected samples, when comparing to the correspondent non-transfected samples. Furthermore, when comparing IP buffer A with IP buffer B, in non-transfected samples and in transfected samples, lysis with IP buffer B seems to lead to a higher amount of Hdm2 on loaded samples.

Thus, considering the aforementioned results, a new IP assay was performed. 80 % confluent A549 cells were transfected with 4 µg of *pΔN6_HDM2*, followed by lysis with IP buffer B 24 h later. Afterwards, we used a new batch of 4B2 monoclonal antibody, which has never been used, to immunoprecipitate Hdm2 isoforms. Unlike the 4B2 aliquot used before (Figure 4.27), this one should not contain denatured antibody with a compromised function, since it never went through repeated freeze/thaw cycles. As we were testing this new antibody aliquot, we did not include the MOCK control. TL, IP and FT samples were collected for Western blot analysis with SMP14 (Figure 4.29).

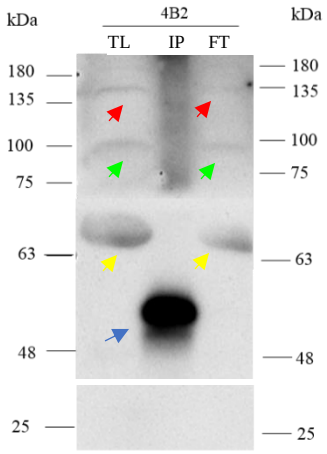


Figure 4. 29 — New 4B2 Hdm2 monoclonal antibody successfully bound to agarose recombinant protein G-containing beads, but no Hdm2 is detected on IP sample. Immunoprecipitation of Hdm2 using 4B2 antibody diluted 1:100 from a cleared lysate of 100 % confluent A549 cells transfected with *pΔN6_HDM2* and lysed with IP buffer B. For immunoprecipitation assay, total lysate of a 100-mm dish was used. Green arrows indicate 90 kDa FL-Hdm2, yellow arrows indicate the isoform or cleavage product with a molecular weight around 60 kDa, red arrows indicate 135 kDa post-translationally modified Hdm2 isoform and blue arrow indicates 4B2 monoclonal antibody heavy chain. TL, total lysate; IP, immunoprecipitated samples; FT, flow through. NzyColour Protein Marker II was used. kDa, kilodalton.

For the first time, we can detect Hdm2 isoforms in TL and FT. Not only FL-Hdm2 is detected (green arrows) but also the isoform or cleavage product with a molecular weight around 60 kDa (yellow arrows). A protein with approximately 135 kDa is also detected (red arrows). It may correspond to a post-translationally modified Hdm2 isoform such as auto ubiquitinated FL-Hdm2.[138] Furthermore, this new aliquot of 4B2 monoclonal antibody seems to contain a functional 4B2, since 4B2 is bound to the beads. Nevertheless, we only see its heavy chain in IP sample (blue arrow). Moreover, no Hdm2 isoforms are detected in IP sample. If the isoform or cleavage product with a molecular weight around 60 kDa does not contain 4B2 epitope, we are not expecting its detection in IP sample. However, FL-Hdm2 is recognized by 4B2; thus, a 90-kDa band should be observed. Then, knowing that FL-p53 interacts with Hdm2 through its N-terminus and knowing that Δ40p53 does not have part of FL-p53 N-terminus, including Hdm2 binding site, and therefore cannot bind to this E3 ubiquitin ligase[139], we stripped off our membrane and incubated it with CM-1 p53 monoclonal antibody. This antibody can detect almost all p53 isoforms, including the FL-p53 and the Δ40p53.[140] (Figure 4.30). FL-p53 but not Δ40p53 detection in IP sample give us an indirect indication that Hdm2 was successfully immunoprecipitated, although we were not able to see it directly.

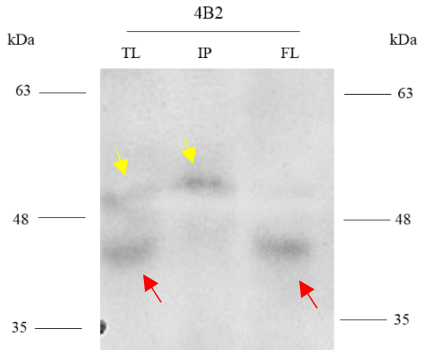


Figure 4. 30 — FL-p53 but not Δ40p53 is detected in IP sample, indicating that Hdm2 was successfully immunoprecipitated. Immunoprecipitation of Hdm2 using 4B2 antibody diluted 1:100 from a cleared lysate of 100 % confluent A549 cells transfected with *pΔN6_HDM2* and lysed with IP buffer B. For immunoprecipitation assay, total lysate of a 100-mm dish was used. CM-1 monoclonal antibody was used to detect p53 isoforms. Yellow arrows indicate FL-p53 and red arrows indicate Δ40p53. TL, total lysate; IP, immunoprecipitated samples; FT, flow through. NzyColour Protein Marker II was used. kDa, kilodalton.

As we can see in Figure 4.30, both FL-p53 (53 kDa; yellow arrow) and Δ40p53 (47 kDa; red arrow) are detected in TL sample; however, only FL-p53 is detected in IP sample. Therefore, Hdm2 was successfully immunoprecipitated by 4B2. Nevertheless, as we could not directly detect Hdm2 in IP sample, this IP assay was not efficient, thus, further optimizations are still required.

Then, to increase Hdm2 levels, and hence augment the input for 4B2 binding, we changed the used of A549 cells for Human Embryonic Kidney 293T cells (HEK 293T), a cell line that allows a more efficient transfection of plasmid DNAs, with a consequent higher expression yield.[141,142] HEK-293T cells are not a cancer cell line; thus, the Hdm2-regulated mRNAs expressed in these cells may not be the same as the ones expressed in A549 cells. This might compromise the further identification of cancer-related mRNAs regulated by Hdm2. However, if we reproduce the stress conditions that characterize tumor microenvironment, we may be able to induce cancer-related mRNAs in this non-tumorigenic cell line and circumvent this issue. Nevertheless, we first need to finish Hdm2 IP optimizations. Thus, besides the change for HEK-293T cell line, we also considered the use of MG132, a proteasome inhibitor, to prevent Hdm2 degradation.[128] Before starting another IP assay, we performed a Western blot analysis of these new conditions. Thus, 80 % confluent HEK-293T cells were transfected with 1.5 µg of *pΔN6_HDM2,* followed by 0.1 % (v/v) DMSO (vehicle) or 25 µM[128] of MG132 (Calbiochem) treatment, 4 h prior to lysis with IP buffer B. A non-transfected control was also included. Afterwards, we evaluated Hdm2 expression in the absence or presence of MG132 by Western blot, using 4B2 monoclonal antibody.

As we can see in Figure 4.31, we only detect a band around 130 kDa (red arrows), which may correspond to autoubiquitinated FL-Hdm2.[138] Moreover, when comparing 130 kDa Hdm2 expression in non-transfected cells with the corresponding transfected cells, no differences are detected. Knowing that Hdm2 overexpression is common in various types of cancer and was already associated with cell motility and invasiveness[143], it is plausible that HEK-293T have mechanisms to maintain this protein at low levels. Then, almost all translated Hdm2 may be being degraded, explaining why only autoubiquitinated Hdm2 is detected. Additionally, when comparing untreated and treated cells with MG132, unexpectedly, no differences in Hdm2 protein levels are detected. Searching for an explanation, we realized that we used an 2013 aliquot. Storage manufacturer's instructions mention that MG132 reconstituted aliquots should be stored only up to 1 month at – 20 ˚C.[144] Therefore, the used aliquot should contain MG132 proteasome inhibitor with its activity compromised, explaining the obtained results.

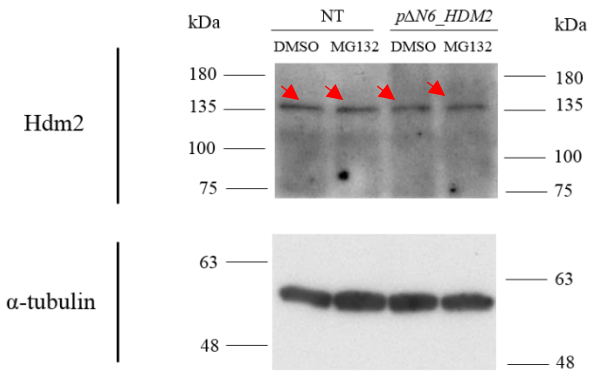


Figure 4. 31 — MG132 treatment of HEK-293T cells transfected with *pΔN6_HDM2* does not increase Hdm2 levels. Western blot of Hdm2 using 4B2 antibody diluted 1:500 from a cleared lysate of 100 % confluent HEK-293T cells transfected with *pΔN6_HDM2*, followed by treatment with either 0.1 % (v/v) DMSO (vehicle) or 25 µM of MG132 (Calbiochem), 4 h prior to lysis with IP buffer B. Red arrows indicate a 135 kDa post-translationally modified Hdm2 isoform. Non-transfected cells were also included as a transfection control. NzyColour Protein Marker II was used. kDa, kilodalton.

The obtained results seem to indicate that almost all translated Hdm2 is being degraded, since only autoubiquitinated Hdm2 is detected. If HEK-293T cells have mechanisms to keep this protein at low cellular levels, its transfection may be accelerating its degradation. Therefore, we tried to immunoprecipate endogenous HEK-293T Hdm2. Then, no transfection was performed. HEK-293T were seeded in 100-mm dishes, followed by lysis with IP buffer B 2 days later. For IP assay, we used 4B2 monoclonal antibody diluted 1:100. A MOCK control was also included. TL, IP and FT samples were harvested for Western blot analysis with 4B11 Hdm2 monoclonal antibody (Figure 4.32). 4B11 epitope spans from amino acid 429 to amino acid 491, within Hdm2 RING domain (Figure 4.26). Therefore, 4B11 can bind to all Hdm2 isoforms capable of interacting with cellular mRNAs, including FL-Hdm2, Hdm2-A, Hdm2-B and Hdm2-C. Therefore, we expected to detect all these isoforms in TL sample. As we used 4B2 for immunoprecipitation, we expected to detect FL-Hdm2 and Hdm2-C in IP sample with 4B2. In IP MOCK sample, we are not expecting the detection of any protein, unless unspecific binding to the beads occurred. In FT samples, we also expected the detection of FL-Hdm2, Hdm2-A, Hdm2-B and Hdm2-C. In FT from 4B2 immunoprecipitation, FL-Hdm2 and Hdm2-C may not appear or may present less protein levels, since the binding of these proteins to 4B2 in IP sample should have occurred.

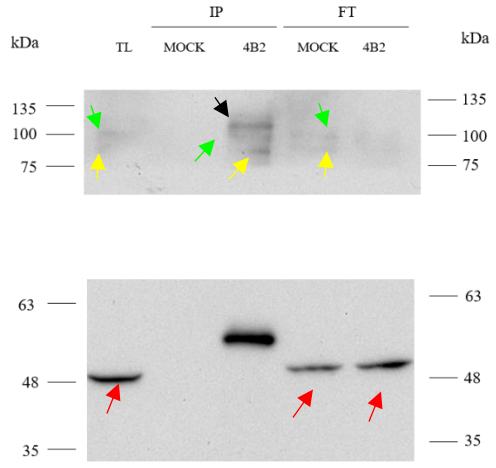


Figure 4. 32 — Hdm2 immunoprecipitation was successfully achieved. Immunoprecipitation of Hdm2 using 4B2 antibody diluted 1:100 from a cleared lysate of 100 % confluent HEK-293T cells lysed with IP buffer B. MOCK control was prepared through incubation of cleared lysate with beads, without any antibody incubation. For each immunoprecipitation, the

total lysate of a 100-mm dish was used. Green arrows indicate 90 kDa FL-Hdm2, yellow arrows indicate 85 kDa Hdm2-C isoform, red arrows indicate 48 kDa Hdm2-B isoform and black arrow indicates a post-translationally modified Hdm2 isoform. TL, total lysate; IP, immunoprecipitated samples; FT, flow through. NzyColour Protein Marker II was used. kDa, kilodalton.

As we can see in Figure 4.32, TL sample shows two diffuse bands between 75 and 100 kDa. This may correspond to FL-Hdm2 (90 kDa; green arrow) and Hdm2-C (85 kDa; yellow arrow). It is also detected an isoform with 48 kDa, possibly Hdm2-B (red arrow). In IP samples, MOCK does not show any band, which means that unspecific ligation to the beads did not occur. In terms of IP sample with 4B2, we see two bands between 75 and 100 kDa, which may also correspond to FL-Hdm2 (green arrow) and Hdm2-C (yellow arrow). It cannot be Hdm2-A, because this isoform does not have 4B2 epitope. We also see a band with more than 100 kDa (black arrow). Having a molecular weight above FL-Hdm2, this unidentified band may correspond to an Hdm2 isoform with post-translational modifications. Nevertheless, it must contain 4B2 epitope. Furthermore, looking for FT samples, both in MOCK and 4B2, Hdm2 isoform with 48 kDa (red arrow), possibly Hdm2-B, is detected. As this isoform cannot bind to 4B2, makes sense that its intensity is similar in both FT samples. Moreover, this seems to be the most expressed Hdm2 isoform in HEK-293T cells. Besides 48 kDa Hdm2 isoform detection in FT from MOCK, two bands between 75 and 100 kDa, possibly FL-Hdm2 (green arrow) and Hdm2-C (yellow arrow), are also identified. These two isoforms do not appear in FT from 4B2 lane. This may indicate that almost all FL-Hdm2 and Hdm2-C bound to 4B2 in IP sample. Thus, considering all these results, Hdm2 isoform immunoprecipitation with 4B2 was successfully achieved.

The next step is RNA extraction from TL, IP and FL samples, followed by reverse transcriptase real-time quantitative PCR analysis. Using both negative and positive mRNA controls for Hdm2 interaction, we will be able to assess the success of this IP assay. We can only proceed to RNA sequencing upon confirming that the system is working properly. It is noteworthy that in our assay we are not immunoprecipitating FL-Hdm2 only. Therefore, Hdm2-interacting mRNAs that will be identified by RNA sequencing could be bound to different Hdm2 isoforms. In fact, most of them contain the RNA-binding RING finger domain. Thus, they may also function as ITAFs in IRES-mediated translation of transcripts related to cellular homeostasis. Nevertheless, knowing which IRES-containing mRNAs each isoform regulates is not the most important. In fact, the most relevant is the unveiling of new mRNAs containing IRES elements regulated by cellular expressed Hdm2 isoforms, in stress conditions similar to those that occur in cancer cells. That will be a significant step to the unveil of cancer mechanisms. Furthermore, it is known that Hdm2 can be regulated by several post-translational modifications that alter not only its function, but also its subcellular localization.[145] For instance, during DNA damage, FL-p53 interaction with Hdm2 is inhibited through Hdm2 phosphorylation on tyrosine 394, which allows p53 tumor suppressor accumulation in the cell.[146] Moreover, after mitogen-induced activation, Akt/PKB serine-threonine kinases phosphorylate Hdm2 on serines 166 and 186, leading to Hdm2 translocation from the cytoplasm to the nucleus.[146] Additionally, during genotoxic stress, ATM phosphorylates Hdm2 at serine 395, favoring *p53* mRNA interaction with Hdm2 and inhibiting p53 protein degradation.[127] Thus, induction of specific post-translational Hdm2 modifications, followed by assessment of which IRES-containing mRNAs that interact with Hdm2 are being translated, would be of great relevance. Additionally, it would also be interesting to find out Hdm2 subcellular localization when Hdm2–IRES-containing mRNA interactions occur. It is known that Hdm2 interaction with *p53* mRNA begins in the nucleus during genotoxic stress, after Hdm2 phosphorylation by ATM and is followed by this protein–RNA complex translocation to the cytoplasm with p53 synthesis induction.[127] However, cellular stress and DNA damage induced by irradiation leads to Hdm2 dephosphorylation and cytoplasmic subcellular localization, with a consequent IRES-mediated translation of *XIAP* mRNA through an Hdm2-dependent manner.[129] But, unlike *p53* mRNA, *XIAP* IRES binds to Hdm2 in the cytoplasm.[129] Thus, Hdm2 seems to regulate IRES elements through different mechanisms and for some

IRES a “nuclear experience” with Hdm2 may be required, whereas for others it may not. Therefore, understanding which cellular stresses trigger a specific mechanism and which IRES-containing mRNAs are induced in each case are questions that are worth of attention as they can provide important knowledge of Hdm2 role as ITAF under different cellular stress conditions.

5. Conclusion and Future Perspectives

At least half of all tumors exhibit mutations in the *TP53* gene.[63] Indeed, the fact that p53 is frequently mutated in cancer led to its identification as an oncogene, when first described in 1979.[147] Later, it was classified as a tumor suppressor, due to its role in maintaining genome integrity and preventing malignant transformation.[62–64]. p53 works as a transcription factor by regulating many biological processes such as cell cycle, apoptosis, stem cell differentiation, senescence and DNA repair.[65] However, unlike many other tumor suppressors, which are frequently inactivated when mutated in tumors, p53 mutations are usually missense mutations that lead to the production of a full-length mutant protein and isoform variants.[63] Besides loss of tumor suppressor activity, tumor-associated mutant p53 proteins often gain new tumorigenic activities that promote angiogenesis, migration, invasion, metastasis, chemoresistance and metabolic changes that favor tumor cell survival.[63] Recently, these tumorigenic functions were associated with Δ160p53 isoform expression. In fact, in a recent study, R273H p53 mutant showed to promote tumor development in a Δ160p53-dependent manner.[75] Moreover, R175H and R248Q along with R273H showed to induce Δ160p53 overexpression, by enhancing Δ160p53 translation.[75] Thus, p53 missense mutations and Δ160p53 overexpression seem to play an important and combined role in cancer development and progression. Furthermore, Δ160p53 expression was recently associated with an IRES element located within the first 432 nucleotides of Δ160p53 coding region.[77] IRES-mediated translation is an alternative translation initiation mechanism that does not require the 5' cap structure.[27] It is frequently used under unfavorable or energy-depriving conditions, such as endoplasmic reticulum (ER) stress, nutrient starvation and mitosis, during which cap-dependent translation is impaired.[27,38] This alternative mechanism of translation initiation, along with others, allows the selective expression of specific mRNAs required to overcome the stress.[27,29] However, tumor cells take advantage of these mechanisms used by normal cells in response to energy-depriving or cellular stress events to cope with unfavorable conditions that characterize tumor microenvironment (for instance, nutrient deprivation, hypoxia, genotoxic and oxidative stress) and proliferate.[27,53] Therefore, the recently identified Δ*160p53* IRES might have an important role in Δ160p53 tumorigenic functions.

Thus, in this master's thesis, we proposed to study the regulation of the expression of alternative protein isoforms involved in carcinogenesis, more specifically, the regulation of Δ160p53 expression through its IRES element, under stress conditions known to impair cap-dependent translation. Using Δ*160p53* IRES-containing bicistronic constructs, we assessed Δ*160p53* IRES activity under normal and ER stress conditions induced by Thapsigargin. In our results, Δ*160p53* 5' UTR showed to have an inhibitory effect in Δ*160p53* IRES activity, in HeLa cells, corroborating the previous results obtained by Marques-Ramos in H1299 and A549 cells[77]. Furthermore, we also assessed Δ*160p53* IRES activity in the presence of R175H, R248Q, R273H and R282W p53 mutations. Only R175H showed to revert some of the Δ*160p53* 5' UTR inhibitory effect in the Δ*160p53* IRES activity, under 2 µM Thapsigargin ER-induced stress. Unexpectedly, R248Q, R273H and R282W missense mutations could not induce Δ*160p53* IRES. Our hypothesis is that HeLa cells may not depend on Δ160p53 expression for tumor cell proliferation, thus Δ160p53 pathways may not be fully active. Indeed, we did not test Δ160p53 endogenous levels in HeLa cells. Moreover, if R175H mutation is the most efficient of all tested mutations in counteracting Δ*160p53* 5' UTR, we may be able to see its effect in HeLa cells, but not the effect of R248Q, R273H and R282W missense mutations. Thus, for the future, we propose the identification and the study of cancer cells lines that highly depend on Δ160p53 expression for survival and proliferation. If p53 missense mutations really induce Δ*160p53* IRES activity to a great extent in those cell lines, therapeutic strategies focusing on Δ*160p53* IRES can be developed to fight those types

of cancer. For instance, antagonists/drugs, such as antisense oligonucleotides, that can disrupt the IRES itself or prevent IRES interactions with the ribosome and with auxiliary factors necessary for their function can be designed.

Additionally, to get more insight on the regulation of Δ160p53 translation through its IRES element, we also aimed the identification of Δ*160p53* IRES *trans*-acting factors, using the MS2 system. This approach takes advantage of *MS2* RNA–MS2 coat protein interaction and requires co-transfection of two plasmids: one containing our sequences of interest upstream twelve *MS2* hairpin repeats and another containing MS2 coat protein sequence.[109–111] Since the latter was already available in the laboratory, we started by cloning p53 sequences of interest upstream *MS2* hairpin repeats. However, we could conclude this phase, since the four tested cloning strategies failed. Nevertheless, a new solution emerged and if it works, co-immunoprecipitation of MS2 coat protein bound to *MS2* hairpin repeats, which in turn are fused to *p53* mRNA sequences, will be done. Afterwards, samples will be sent to mass spectrometry for *p53* mRNA-interacting proteins identification. It is known that many ITAFs are shared between IRES-containing mRNAs and many mRNAs containing IRES elements are frequently overexpressed in cancer cells. [29,45,82,116,117] Thus, deregulation of ITAFs shared between cancer-related mRNAs might play a key role in cancer development, and because of that, the uncover of new *p53* ITAFs, including factors that regulate Δ160p53 oncogenic protein expression, is a matter of great importance.

Finally, since IRES-mediated translation seems to play a key role in cancer development, we also aimed to identify new IRES-containing mRNAs. It is known that, besides Hdm2 role on p53 degradation, this protein can also directly interact with some mRNAs, such *p53* and *XIAP*, functioning as an ITAF and regulating their expression.[128–130] Since these proteins are frequently deregulated in cancer, we hypothesize that Hdm2 may also regulate IRES-mediated translation of other cellular mRNAs commonly involved in tumorigenesis. Then, experimentally, we intended to perform co-immunoprecipitation of Hdm2 with its bound mRNAs, using Hdm2 antibodies, followed by RNA deep sequencing analysis. In this master's thesis, we could only finish Hdm2 immunoprecipitation assay optimizations; assessment of Hdm2-bound RNAs co-immunoprecipitation success still needs to be done, before RNA sequencing. Furthermore, knowing that Hdm2 can be regulated by specific post-translational modifications depending on the stimulus, several cellular stresses can be induced to unveil the selective expression of IRES-containing mRNAs regulated by Hdm2 isoforms. Ultimately, information obtained by RNA sequencing of Hdm2-bound mRNAs and by mass spectrometry of *p53* mRNA-bound ITAFs can be combined. The new IRES-containing mRNAs uncovered by RNA sequencing may also be regulated by p53 newly identified ITAFs. Therefore, immunoprecipitation of newly found ITAFs may be done, followed by RT–qPCR to detect the presence of the novel identified IRES-containing mRNAs.

At the end, concluding all these lines of research, we hope to unveil new understandings regarding IRES-mediated translation of cancer-related mRNAs. Knowing that, when deregulated, this mechanism of translation initiation can promote tumorigenesis, allowing cancer cells to survive, the understanding of how IRES-containing mRNAs are regulated under different stress conditions and how the switch between cell homeostasis and cell neoplastic transformation is triggered, will provide important knowledge for the development of new therapeutic strategies. In fact, cancer cells frequently develop resistance to the conventional therapies, particularly in more advanced stages of cancer progression.[148] Thus, new approaches are required to fight this second leading cause of dead worldwide.[149]

6. References

(1) Orphanides, G.; Reinberg, D. A Unified Theory of Gene Expression. *Cell* **2002**, *108* (4), 439–451.

(2) Maniatis, T.; Reed, R. An Extensive Network of Coupling among Gene Expression Machines. *Nature* **2002**, *416* (6880), 499–506.

(3) Division of Advanced Education in Sciences - The University of Tokyo. Regulation of Gene Expression in Eukaryotic Cells http://csls-text.c.u-tokyo.ac.jp/active/04_03.html (accessed Jun 16, 2018).

(4) Sainsbury, S.; Bernecky, C.; Cramer, P. Structural Basis of Transcription Initiation by RNA Polymerase II. *Nat. Rev. Mol. Cell Biol.* **2015**, *16* (3), 129–143.

(5) Luna, R.; Gaillard, H.; González-Aguilera, C.; Aguilera, A. Biogenesis of mRNPs: Integrating Different Processes in the Eukaryotic Nucleus. *Chromosoma* **2008**, *117* (4), 319–331.

(6) Ficket, J. W.; Hatzigeorgiou, A. G. Eukaryotic Promoter Recognition. *Genome Res.* **1997**, *7* (9), 861–878.

(7) Proudfoot, N. J.; Furger, A.; Dye, M. J. Integrating mRNA Processing with Transcription. *Cell* **2002**, *108* (4), 501–512.

(8) Richard, P.; Manley, J. L. Transcription Termination by Nuclear RNA Polymerases. *Genes Dev.* **2009**, *23* (11), 1247–1269.

(9) Hocine, S.; Singer, R. H.; Grünwald, D. RNA Processing and Export. *Cold Spring Harb. Perspect. Biol.* **2010**, *2* (12), a000752.

(10) Shi, Y.; Manley, J. L. The End of the Message: Multiple Protein-RNA Interactions Define the mRNA Polyadenylation Site. *Genes Dev.* **2015**, *29* (9), 889–897.

(11) Iglesias, N.; Stutz, F. Regulation of mRNP Dynamics along the Export Pathway. *FEBS Lett.* **2008**, *582* (14), 1987–1996.

(12) Gebauer, F.; Hentze, M. W. Molecular Mechanisms of Translational Control. *Nat. Rev. Mol. Cell Biol.* **2004**, *5* (10), 827–835.

(13) Preiss, T.; Hentze, M. W. Starting the Protein Synthesis Machine: Eukaryotic Translation Initiation. *BioEssays* **2003**, *25* (12), 1201–1211.

(14) Ramakrishnan, V. Ribosome Structure and the Mechanism of Translation. *Cell* **2002**, *108* (4), 557–572.

(15) Sonenberg, N.; Hinnebusch, A. G. Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell* **2009**, *136* (4), 731–745.

(16) Haimov, O.; Sinvani, H.; Dikstein, R. Cap-Dependent, Scanning-Free Translation Initiation Mechanisms. *Biochim. Biophys. Acta - Gene Regul. Mech.* **2015**, *1849* (11), 1313–1318.

(17) Hinnebusch, A. G. The Scanning Mechanism of Eukaryotic Translation Initiation. *Annu. Rev. Biochem.* **2014**, *83* (1), 779–812.

(18) Preiss, T.; Hentze, M. W. Starting the Protein Synthesis Machine: Eukaryotic Translation Initiation. *BioEssays* **2003**, *25* (12), 1201–1211.

(19) Aitken, C. E.; Lorsch, J. R. A Mechanistic Overview of Translation Initiation in Eukaryotes. *Nat. Struct. Mol. Biol.* **2012**, *19* (6), 568–576.

(20) Kapp, L. D.; Lorsch, J. R. The Molecular Mechanics of Eukaryotic Translation. *Annu. Rev. Biochem.* **2004**, *73* (1), 657–704.

(21) Andreou, A. Z.; Harms, U.; Klostermeier, D. eIF4B Stimulates eIF4A ATPase and Unwinding Activities by Direct Interaction through Its 7-Repeats Region. *RNA Biol.* **2017**, *14* (1), 113–123.

(22) Korostelev, A. A. A Deeper Look into Translation Initiation. *Cell* **2014**, *159* (3), 475–476.

(23) Dever, T. E.; Green, R. The Elongation, Termination, and Recycling Phases of Translation in Eukaryotes. *Cold Spring Harb. Perspect. Biol.* **2012**, *4* (7), a013706.

(24) Agirrezabala, X.; Frank, J. Elongation in Translation as a Dynamic Interaction among the Ribosome, tRNA, and Elongation Factors EF-G and EF-Tu. *Q. Rev. Biophys.* **2009**, *42* (3), 159–200.

(25) Salas-Marco, J.; Bedwell, D. M. GTP Hydrolysis by eRF3 Facilitates Stop Codon Decoding during Eukaryotic Translation Termination. *Mol. Cell. Biol.* **2004**, *24* (17), 7769–7778.

(26) Yamamoto, H.; Unbehaun, A.; Spahn, C. M. T. Ribosomal Chamber Music: Toward an Understanding of IRES Mechanisms. *Trends Biochem. Sci.* **2017**, *42* (8), 655–668.

(27) Lacerda, R.; Menezes, J.; Romão, L. More than Just Scanning: The Importance of Cap-Independent mRNA Translation Initiation for Cellular Stress Response and Cancer. *Cell. Mol. Life Sci.* **2017**, *74* (9), 1659–1680.

(28) Komar, A. A.; Hatzoglou, M. Internal Ribosome Entry Sites in Cellular mRNAs: Mystery of Their Existence. *J. Biol. Chem.* **2005**, *280* (25), 23425–23428.

(29) Komar, A. A.; Hatzoglou, M. Cellular IRES-Mediated Translation: The War of ITAFs in Pathophysiological States. *Cell Cycle* **2011**, *10* (2), 229–240.

(30) Baird, S. D. Searching for IRES. *RNA* **2006**, *12* (10), 1755–1785.

(31) Chard, L. S.; Kaku, Y.; Jones, B.; Nayak, A.; Belsham, G. J. Functional Analyses of RNA Structures Shared between the Internal Ribosome Entry Sites of Hepatitis C Virus and the Picornavirus Porcine Teschovirus 1 Talfan. *J. Virol.* **2006**, *80* (3), 1271–1279.

(32) Filbin, M. E.; Kieft, J. S. Toward a Structural Understanding of IRES RNA Function. *Current Opinion in Structural Biology*. 2009, pp 267–276.

(33) Kieft, J. S. Viral IRES RNA Structures and Ribosome Interactions. *Trends in Biochemical Sciences*. 2008, pp 274–283.

(34) Gritsenko, A. A.; Weingarten-Gabbay, S.; Elias-Kirma, S.; Nir, R.; de Ridder, D.; Segal, E. Sequence Features of Viral and Human Internal Ribosome Entry Sites Predictive of Their Activity. *PLoS Comput. Biol.* **2017**, *13* (9), e1005734.

(35) Weingarten-Gabbay, S.; Elias-Kirma, S.; Nir, R.; Gritsenko, A. A.; Stern-Ginossar, N.; Yakhini, Z.; Weinberger, A.; Segal, E. Comparative Genetics: Systematic Discovery of Cap-Independent Translation Sequences in Human and Viral Genomes. *Science (80-.).* **2016**, *351* (6270), aad4939.

(36) Okuda, T.; Nishimura, M.; Nakao, M.; Fujita, Y. RUNX1/AML1: A Central Player in Hematopoiesis. *Int. J. Hematol.* **2001**, *74* (3), 252–257.

(37) Leppek, K.; Das, R.; Barna, M. Functional 5′ UTR mRNA Structures in Eukaryotic Translation Regulation and How to Find Them. *Nat. Rev. Mol. Cell Biol.* **2018**, *19* (3), 158–174.

(38) Liu, B.; Qian, S. B. Translational Reprogramming in Cellular Stress Response. *Wiley Interdiscip. Rev. RNA* **2014**, *5* (3), 301–305.

(39) Showkat, M.; Beigh, M. A.; Andrabi, K. I. mTOR Signaling in Protein Translation Regulation: Implications in Cancer Genesis and Therapeutic Interventions. *Mol. Biol. Int.* **2014**, *2014*, 1–14.

(40) Nandagopal, N.; Roux, P. P. Regulation of Global and Specific mRNA Translation by the mTOR Signaling Pathway. *Transl.* **2015**, *3* (1), e983402.

(41) Baird, T. D.; Wek, R. C. Eukaryotic Initiation Factor 2 Phosphorylation and Translational Control in Metabolism. *Adv. Nutr. An Int. Rev. J.* **2012**, *3* (3), 307–321.

(42) Pinkstaff, J. K.; Chappell, S. A.; Mauro, V. P.; Edelman, G. M.; Krushel, L. A. Internal Initiation of Translation of Five Dendritically Localized Neuronal mRNAs. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (5), 2770–2775.

(43) King, H. A.; Cobbold, L. C.; Willis, A. E. The Role of IRES Trans -Acting Factors in Regulating Translation Initiation. *Biochem. Soc. Trans.* **2010**, *38* (6), 1581–1586.

(44) Bushell, Martin, Stoneley, Mark, Kong, Yi Wen, Hamilton, Tiffany L., Spriggs, Keith A., , Dobbyn, Helen C., Qin, Xiaoli, Sarnow, Peter, Willis, A. E. Polypyrimidine Tract Binding Protein Regulates IRES-Mediated Gene Expression during Apoptosis. *Mol. Cell* **2006**, *23* (3), 401–412.

(45) Stoneley, M.; Willis, A. E. Cellular Internal Ribosome Entry Segments: Structures, Trans-Acting Factors and Regulation of Gene Expression. *Oncogene* **2004**, *23* (18), 3200–3207.

(46) Evans, J. R.; Mitchell, S. A.; Spriggs, K. A.; Ostrowski, J.; Bomsztyk, K.; Ostarek, D.; Willis, A. E. Members of the Poly (rC) Binding Protein Family Stimulate the Activity of the c-Myc Internal Ribosome Entry Segment in Vitro and in Vivo. *Oncogene* **2003**, *22* (39), 8012–8020.

(47) Nesbit, C. E.; Grove, L. E.; Yin, X.; Prochownik, E. V. Differential Apoptotic Behaviors of c-Myc, N-Myc, and L-Myc Oncoproteins. *Cell Growth Dfferentiation* **1998**, *9* (9), 731–742.

(48) Kim, Y. K.; Hahm, B.; Jang, S. K. Polypyrimidine Tract-Binding Protein Inhibits Translation of Bip mRNA. *J. Mol. Biol.* **2000**, *304* (2), 119–133.

(49) Spriggs, K. A.; Cobbold, L. C.; Jopling, C. L.; Cooper, R. E.; Wilson, L. A.; Stoneley, M.; Coldwell, M. J.; Poncet, D.; Shen, Y.-C.; Morley, S. J.; et al. Canonical Initiation Factor Requirements of the Myc Family of Internal Ribosome Entry Segments. *Mol. Cell. Biol.* **2009**, *29* (6), 1565–1574.

(50) Thakor, N.; Holcik, M. IRES-Mediated Translation of Cellular Messenger RNA Operates in eIF2α-Independent Manner during Stress. *Nucleic Acids Res.* **2012**, *40* (2), 541–552.

(51) Holcik, M. Could the eIF2α-Independent Translation Be the Achilles Heel of Cancer? *Front. Oncol.* **2015**, *5* (264).

(52) Ozretić, P.; Bisio, A.; Inga, A.; Levanat, S. The Growing Relevance of Cap-Independent Translation Initiation in Cancer-Related Genes. *Period. Biol.* **2012**, *114* (4), 471–478.

(53) Walters, B.; Thompson, S. R. Cap-Independent Translational Control of Carcinogenesis. *Front. Oncol.* **2016**, *6* (128), 1–8.

(54) Piret, J. P.; Mottet, D.; Raes, M.; Michiels, C. Is HIF-1α a pro- or an Anti-Apoptotic Protein? *Biochem. Pharmacol.* **2002**, *64* (5-6), 889–892.

(55) Ramakrishnan, S.; Anand, V.; Roy, S. Vascular Endothelial Growth Factor Signaling in Hypoxia and Inflammation. *J. neuroimmune Pharmacol.* **2014**, *9* (2), 142–160.

(56) Holcík, M. Targeting Translation for Treatment of Cancer--a Novel Role for IRES? *Curr. Cancer Drug Targets* **2004**, *4* (3), 299–311.

(57) Faye, M. D.; Holcik, M. The Role of IRES Trans-Acting Factors in Carcinogenesis. *Biochim. Biophys. Acta - Gene Regul. Mech.* **2015**, *1849* (7), 887–897.

(58) Komar, A. A.; Hatzoglou, M. Exploring Internal Ribosome Entry Sites as Therapeutic Targets.

Front. Oncol. **2015**, *5*, 233.

(59) Martinand-Mari, C.; Lebleu, B.; Robbins, I. Oligonucleotide-Based Strategies to Inhibit Human Hepatitis C Virus. *Oligonucleotides* **2003**, *13* (6), 539–548.

(60) Renaud-Gabardos, E. Internal Ribosome Entry Site-Based Vectors for Combined Gene Therapy. *World J. Exp. Med.* **2015**, *5* (1), 11–20.

(61) Rayssac, A.; Neveu, C.; Pucelle, M.; Van den Berghe, L.; Prado-Lourenco, L.; Arnal, J.-F.; Chaufour, X.; Prats, A.-C. IRES-Based Vector Coexpressing FGF2 and Cyr61 Provides Synergistic and Safe Therapeutics of Lower Limb Ischemia. *Mol. Ther.* **2009**, *17* (12), 2010–2019.

(62) Ji, B.; Harris, B. R. E.; Liu, Y.; Deng, Y.; Gradilone, S. A.; Cleary, M. P.; Liu, J.; Yang, D. Q. Targeting IRES-Mediated p53 Synthesis for Cancer Diagnosis and Therapeutics. *Int. J. Mol. Sci.* **2017**, *18* (1), 93.

(63) Liu, J.; Zhang, C.; Hu, W.; Feng, Z. Tumor Suppressor p53 and Its Mutants in Cancer Metabolism. *Cancer Lett.* **2015**, *356* (2), 197–203.

(64) Sharathchandra, A.; Katoch, A.; Das, S. IRES Mediated Translational Regulation of p53 Isoforms. *Wiley Interdiscip. Rev. RNA* **2014**, *5* (1), 131–139.

(65) Vousden, K. H.; Prives, C. Blinded by the Light: The Growing Complexity of p53. *Cell* **2009**, *137* (3), 413–431.

(66) Meek, D. W. Regulation of the p53 Response and Its Relationship to Cancer. *Biochem. J.* **2015**, *469* (3), 325–346.

(67) Zilfou, J. T.; Lowe, S. W. Tumor Suppressive Functions of p53. *Cold Spring Harb. Perspect. Biol.* **2009**, *1* (5), a001883–a001883.

(68) Candeias, M. M. The Can and Can't Dos of p53 RNA. *Biochimie* **2011**, *93* (11), 1962–1965.

(69) Surget, S.; Khoury, M. P.; Bourdon, J. C. Uncovering the Role of p53 Splice Variants in Human Malignancy: A Clinical Perspective. *Onco. Targets. Ther.* **2013**, *7*, 57–67.

(70) Senturk, S.; Yao, Z.; Camiolo, M.; Stiles, B.; Rathod, T.; Walsh, A. M.; Nemajerova, A.; Lazzara, M. J.; Altorki, N. K.; Krainer, A.; et al. p53ψ Is a Transcriptionally Inactive p53 Isoform Able to Reprogram Cells toward a Metastatic-like State. *Proc. Natl. Acad. Sci.* **2014**, *111* (32), E3287–E3296.

(71) Rohaly, G.; Chemnitz, J.; Dehde, S.; Nunez, A. M.; Heukeshoven, J.; Deppert, W.; Dornreiter, I. A Novel Human p53 Isoform Is an Essential Element of the ATR-Intra-S Phase Checkpoint. *Cell* **2005**, *122* (1), 21–32.

(72) Pekova, S.; Cmejla, R.; Smolej, L.; Kozak, T.; Spacek, M.; Prucha, M. Identification of a Novel, Transactivation-Defective Splicing Variant of p53 Gene in Patients with Chronic Lymphocytic Leukemia. *Leuk. Res.* **2008**, *32* (3), 395–400.

(73) Bourdon, J. C. p53 and Its Isoforms in Cancer. *Br. J. Cancer* **2007**, *97* (3), 277–282.

(74) Khoury, M. P.; Bourdon, J. C. P53 Isoforms: An Intracellular Microprocessor? *Genes and Cancer* **2011**, *2* (4), 453–465.

(75) Candeias, M. M.; Hagiwara, M.; Matsuda, M. Cancer Specific Mutations in p53 Induce the Translation of Δ160p53 Promoting Tumorigenesis. *EMBO Rep.* **2016**, *17* (11), 1542–1551.

(76) Halaby, M. J.; Yang, D. Q. p53 Translational Control: A New Facet of p53 Regulation and Its Implication for Tumorigenesis and Cancer Therapeutics. *Gene* **2007**, *395* (1-2), 1–7.

(77) Marques-Ramos, A. L. Translational Regulation Mediated by Internal Ribosome Entry Sites of

the MTOR and Δ133P53 Human Transcripts, Universidade de Coimbra, 2013.

(78) Lang, G. A.; Iwakuma, T.; Suh, Y. A.; Liu, G.; Rao, V. A.; Parant, J. M.; Valentin-Vega, Y. A.; Terzian, T.; Caldwell, L. C.; Strong, L. C.; et al. Gain of Function of a p53 Hot Spot Mutation in a Mouse Model of Li-Fraumeni Syndrome. *Cell* **2004**, *119* (6), 861–872.

(79) Olive, K. P.; Tuveson, D. A.; Ruhe, Z. C.; Yin, B.; Willis, N. A.; Bronson, R. T.; Crowley, D.; Jacks, T. Mutant p53 Gain of Function in Two Mouse Models of Li-Fraumeni Syndrome. *Cell* **2004**, *119* (6), 847–860.

(80) Muller, P. A. J.; Vousden, K. H. Mutant p53 in Cancer: New Functions and Therapeutic Opportunities. *Cancer Cell* **2014**, *25* (3), 304–317.

(81) Aschauer, L.; Muller, P. A. J. Novel Targets and Interaction Partners of Mutant p53 Gain-Of-Function. *Biochem. Soc. Trans.* **2016**, *44* (2), 460–466.

(82) Wei, J.; Zaika, E.; Zaika, A. P53 Family: Role of Protein Isoforms in Human Cancer. *J. Nucleic Acids* **2012**, *2012* (19).

(83) Lacerda, R. Non-Canonical Translation Initiation of Proteins with Potential Relevance in Colorectal Cancer, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, 2016.

(84) Neves, A. R. R. Indentification and Characterization of Internal Risobome Entry Sites (IRES) in Cancer Pathways, Faculdade de Ciências da Universidade de Lisboa, 2018.

(85) Wood, K. V. The Chemistry of Bioluminescent Reporter Assays. *Promega Notes* **1998**, *65*, 14–20.

(86) Ozretić, P.; Bisio, A.; Musani, V.; Trnski, D.; Sabol, M.; Levanat, S.; Inga, A. Regulation of Human PTCH1b Expression by Different 5' Untranslated Region Cis-Regulatory Elements. *RNA Biol.* **2015**, *12* (3), 290–304.

(87) Szegezdi, E.; Logue, S. E.; Gorman, A. M.; Samali, A. Mediators of Endoplasmic Reticulum Stress-Induced Apoptosis. *EMBO Rep.* **2006**, *7* (9), 880–885.

(88) Oren, M.; Rotter, V. Mutant p53 Gain-of-Function in Cancer. *Cold Spring Harb. Perspect. Biol.* **2010**, *2* (2), a001107.

(89) Zhang, Y.; Coillie, S. V; Fang, J.-Y.; Xu, J. Gain of Function of Mutant p53: R282W on the Peak? *Oncogenesis* **2016**, *5*, e196.

(90) Baugh, E. H.; Ke, H.; Levine, A. J.; Bonneau, R. A.; Chan, C. S. Why Are There Hotspot Mutations in the TP53 Gene in Human Cancers? *Cell Death Differ.* **2017**, *25* (1), 154–160.

(91) Lu, X. P53: A Target and a Biomarker of Cancer Therapy? In *Recent Advances in Cancer Research and Therapy*; Elsevier, 2012; pp 197–213.

(92) Tan, Y.; Luo, R. Structural and Functional Implications of p53 Missense Cancer Mutations. *PMC Biophys.* **2009**, *2* (1), 5.

(93) Chappell, S. A.; Dresios, J.; Edelman, G. M.; Mauro, V. P. Ribosomal Shunting Mediated by a Translational Enhancer Element That Base Pairs to 18S rRNA. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (25), 9488–9493.

(94) Terenin, I. M.; Smirnova, V. V.; Andreev, D. E.; Dmitriev, S. E.; Shatsky, I. N. A Researcher's Guide to the Galaxy of IRESs. *Cell. Mol. Life Sci.* **2017**, *74* (8), 1431–1455.

(95) The Human Protein Atlas. GAPDH https://www.proteinatlas.org/ENSG00000111640-GAPDH/cell#rna (accessed Sep 19, 2018).

(96) NZYTech Lda. - Genes and Enzymes. NZY Reverse Transcriptase https://www.nzytech.com/products-services/reverse-transcriptases/mb124/ (accessed Aug 3,

2018).

(97) Thermo Fisher Scientific. SuperScript II Reverse Transcriptase https://www.thermofisher.com/order/catalog/product/18064014 (accessed Aug 3, 2018).

(98) Michael L. Altshuler. PCR Troubleshooting: The Template DNA. In *PCR Troubleshooting: The Essential Guide*; Caister Academic Press: Wymondham, 2006; p 18.

(99) Rychlik, W.; Spencer, W. J.; Rhoads, R. E. Optimization of the Annealing Temperature for DNA Amplification in Vitro. *Nucleic Acids Res.* **1991**, *19* (3), 6409–6412.

(100) Ellsworth, D. L.; Rittenhouse, K. D.; Honeycutt, R. L. Artifactual Variation in Randomly Amplified Polymorphic DNA Banding Patterns. *Biotechniques* **1993**, *14* (2), 214–217.

(101) Williams, J. F. Optimization Strategies for the Polymerase Chain Reaction. *Biotechniques* **1989**, *7* (7), 762–769.

(102) Eckert, K. A.; Kunkel, T. A. High Fidelity DNA Synthesis by the Thermus Aquaticus DNA Polymerase. *Nucleic Acids Res.* **1990**, *18* (13), 3739–3744.

(103) Jackson, R. J. The Current Status of Vertebrate Cellular mRNA IRESs. *Cold Spring Harb. Perspect. Biol.* **2013**, *5* (2), a011569.

(104) Holcík, M.; Gordon, B. W.; Korneluk, R. G. The Internal Ribosome Entry Site-Mediated Translation of Antiapoptotic Protein XIAP Is Modulated by the Heterogeneous Nuclear Ribonucleoproteins C1 and C2. *Mol. Cell. Biol.* **2003**, *23* (1), 280–288.

(105) Stoneley, M. Analysis of the c-Myc IRES; a Potential Role for Cell-Type Specific Trans-Acting Factors and the Nuclear Compartment. *Nucleic Acids Res.* **2000**, *28* (3), 687–694.

(106) Kozak, M. A Second Look at Cellular mRNA Sequences Said to Function as Internal Ribosome Entry Sites. *Nucleic Acids Res.* **2005**, *33* (20), 6593–6602.

(107) Arita, M.; Zhong, X.; Min, Z.; Hemmi, H.; Shimatake, H. Multiple Sites Required for Expression in 5′-Flanking Region of the hMLH1 Gene. *Gene* **2003**, *306* (1-2), 57–65.

(108) Ito, E.; Yanagisawa, Y.; Iwahashi, Y.; Suzuki, Y.; Nagasaki, H.; Akiyama, Y.; Sugano, S.; Yuasa, Y.; Maruyama, K. A Core Promoter and a Frequent Single-Nucleotide Polymorphism of the Mismatch Repair Gene hMLH1. *Biochem. Biophys. Res. Commun.* **1999**, *256* (3), 488–494.

(109) Johansson, H. E.; Liljas, L.; Uhlenbeck, O. C. RNA Recognition by the MS2 Phage Coat Protein. *Semin. Virol.* **1997**, *8* (3), 176–185.

(110) Slobodin, B.; Gerst, J. E. A Novel mRNA Affinity Purification Technique for the Identification of Interacting Proteins and Transcripts in Ribonucleoprotein Complexes. *RNA* **2010**, *16* (11), 2277–2290.

(111) Yoon, J. H.; Srikantan, S.; Gorospe, M. MS2-TRAP (MS2-Tagged RNA Affinity Purification): Tagging RNA to Identify Associated miRNAs. *Methods* **2012**, *58* (2), 81–87.

(112) Yoon, J. H.; Gorospe, M. Identification of mRNA-Interacting Factors by MS2-TRAP (MS2-Tagged RNA Affinity Purification). *Methods Mol. Biol.* **2016**, *1421*, 15–22.

(113) Chao, J. A.; Czaplinski, K.; Singer, R. H. Using the Bacteriophage MS2 Coat Protein-RNA Binding Interaction to Visualize RNA in Living Cells. In *Probes and Tags to Study Biomolecular Function: for Proteins, RNA, and Membranes*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2008; pp 163–174.

(114) Candeias, M. M.; Powell, D. J.; Roubalova, E.; Apcher, S.; Bourougaa, K.; Vojtesek, B.; Bruzzoni-Giovanelli, H.; Fåhraeus, R. Expression of p53 and p53/47 Are Controlled by Alternative Mechanisms of Messenger RNA Translation Initiation. *Oncogene* **2006**, *25* (52),

6936–6947.

(115) New England Biolabs. Cleavage Close to the End of DNA Fragments https://international.neb.com/tools-and-resources/usage-guidelines/cleavage-close-to-the-end-of-dna-fragments (accessed Aug 9, 2018).

(116) Lu, Y.; Wang, W.; Wang, J.; Yang, C.; Mao, H.; Fu, X.; Wu, Y.; Cai, J.; Han, J.; Xu, Z.; et al. Overexpression of Arginine Transporter CAT-1 Is Associated with Accumulation of L-Arginine and Cell Growth in Human Colorectal Cancer Tissue. *PLoS One* **2013**, *8* (9), e73866.

(117) Cutress, R. I.; Townsend, P. A.; Brimmell, M.; Bateman, A. C.; Hague, A.; Packham, G. Bag-1 Expression and Function in Human Cancer. *Br. J. Cancer* **2002**, *87* (8), 834–839.

(118) King, H. A.; Cobbold, L. C.; Willis, A. E. The Role of IRES Trans -Acting Factors in Regulating Translation Initiation. *Biochem. Soc. Trans.* **2010**, *38* (6), 1581–1586.

(119) Schepens, B.; Tinton, S. A.; Bruynooghe, Y.; Parthoens, E.; Haegman, M.; Beyaert, R.; Cornelis, S. A Role for hnRNP C1/C2 and Unr in Internal Initiation of Translation during Mitosis. *EMBO J.* **2007**, *26* (1), 158–169.

(120) Hussain, A. R.; Siraj, A. K.; Ahmed, M.; Bu, R.; Pratheeshkumar, P.; Alrashed, A. M.; Qadri, Z.; Ajarim, D.; Al-Dayel, F.; Beg, S.; et al. XIAP over-Expression Is an Independent Poor Prognostic Marker in Middle Eastern Breast Cancer and Can Be Targeted to Induce Efficient Apoptosis. *BMC Cancer* **2017**, *17* (1), 640–653.

(121) Miller, D. M.; Thomas, S. D.; Islam, A.; Muench, D.; Sedoris, K. C-Myc and Cancer Metabolism. *Clin. Cancer Res.* **2012**, *18* (20), 5546–5553.

(122) Zhou, Y.; Han, C.; Li, D.; Yu, Z.; Li, F.; Li, F.; An, Q.; Bai, H.; Zhang, X.; Duan, Z.; et al. Cyclin-Dependent Kinase 11p110 (CDK11p110) Is Crucial for Human Breast Cancer Cell Proliferation and Growth. *Sci. Rep.* **2015**, *5* (1), 10433.

(123) Anichini, A.; Mortarini, R.; Sensi, M.; Zanon, M. APAF-1 Signaling in Human Melanoma. *Cancer Lett.* **2006**, *238* (2), 168–179.

(124) Sharma, D. K.; Bressler, K.; Patel, H.; Balasingam, N.; Thakor, N. Role of Eukaryotic Initiation Factors during Cellular Stress and Cancer Progression. *J. Nucleic Acids* **2016**, *2016*, 8235121.

(125) Iwakuma, T.; Lozano, G. MDM2, An Introduction. *Mol. Cancer Res.* **2003**, *1* (14), 993–1000.

(126) Reza Saadatzadeh, M.; Elmi, A. N.; Pandya, P. H.; Bijangi-Vishehsaraei, K.; Ding, J.; Stamatkin, C. W.; Cohen-Gadol, A. A.; Pollok, K. E. The Role of MDM2 in Promoting Genome Stability versus Instability. *Int. J. Mol. Sci.* **2017**, *18* (10), E2216.

(127) Gajjar, M.; Candeias, M. M.; Malbert-Colas, L.; Mazars, A.; Fujita, J.; Olivares-Illana, V.; Fåhraeus, R. The p53 mRNA-Mdm2 Interaction Controls mdm2 Nuclear Trafficking and Is Required for p53 Activation Following Dna Damage. *Cancer Cell* **2012**, *21* (1), 25–35.

(128) Candeias, M. M.; Malbert-Colas, L.; Powell, D. J.; Daskalogianni, C.; Maslon, M. M.; Naski, N.; Bourougaa, K.; Calvo, F.; Fåhraeus, R. p53 mRNA Controls p53 Activity by Managing Mdm2 Functions. *Nat. Cell Biol.* **2008**, *10* (9), 1098–1105.

(129) Gu, L.; Zhu, N.; Zhang, H.; Durden, D. L.; Feng, Y.; Zhou, M. Regulation of XIAP Translation and Induction by MDM2 Following Irradiation. *Cancer Cell* **2009**, *15* (5), 363–375.

(130) Bohlman, S.; Manfredi, J. J. Mdm2-RNA Interactions as a Target for Cancer Therapy: It's Not All About p53. *Cancer Cell* **2016**, *30* (4), 513–514.

(131) Cheng, Q.; Chen, J. The Phenotype of MDM2 Auto-Degradation after DNA Damage Is due to Epitope Masking by Phosphorylation. *Cell Cycle* **2011**, *10* (7), 1162–1166.

(132) Lal, A.; Haynes, S. R.; Gorospe, M. Clean Western Blot Signals from Immunoprecipitated Samples. *Mol. Cell. Probes* **2005**, *19* (6), 385–388.

(133) Abcam. Antibody storage guide https://www.abcam.com/protocols/antibody-storage-guide (accessed Aug 11, 2018).

(134) Vaughan, C.; Mohanraj, L.; Singh, S.; Dumur, C. I.; Ramamoorthy, M.; Garrett, C. T.; Windle, B.; Yeudall, W. A.; Deb, S.; Deb, S. P. Human Oncoprotein MDM2 Up-Regulates Expression of NF-κB2 Precursor p100 Conferring a Survival Advantage to Lung Cells. *Genes and Cancer* **2011**, *2* (10), 943–955.

(135) Vagenende, V.; Yap, M. G. S.; Trout, B. L. Mechanisms of Protein Stabilization and Prevention of Protein Aggregation by Glycerol. *Biochemistry* **2009**, *48* (46), 11084–11096.

(136) Abcam. Sample preparation for western blot https://www.abcam.com/protocols/sample-preparation-for-western-blot (accessed Jul 28, 2018).

(137) Oliver, T. G.; Meylan, E.; Chang, G. P.; Xue, W.; Burke, J. R.; Humpton, T. J.; Hubbard, D.; Bhutkar, A.; Jacks, T. Caspase-2-Mediated Cleavage of Mdm2 Creates a p53-Induced Positive Feedback Loop. *Mol. Cell* **2011**, *43* (1), 57–71.

(138) Sosin, A. M.; Burger, A. M.; Siddiqi, A.; Abrams, J.; Mohammad, R. M.; Al-Katib, A. M. HDM2 Antagonist MI-219 (spiro-Oxindole), but Not Nutlin-3 (cis-Imidazoline), Regulates p53 through Enhanced HDM2 Autoubiquitination and Degradation in Human Malignant B-Cell Lymphomas. *J. Hematol. Oncol.* **2012**, *5*, 57.

(139) Ghosh, A.; Stewart, D.; Matlashewski, G. Regulation of Human p53 Activity and Cell Localization by Alternative Splicing. *Mol. Cell. Biol.* **2004**, *24* (18), 7987–7997.

(140) Khoury, M. P.; Bourdon, J. C. The Isoforms of the p53 Protein. *Cold Spring Harb. Perspect. Biol.* **2010**, *2* (3), a000927.

(141) Vink, T.; Oudshoorn-Dickmann, M.; Roza, M.; Reitsma, J. J.; de Jong, R. N. A Simple, Robust and Highly Efficient Transient Expression System for Producing Antibodies. *Methods* **2014**, *65* (1), 5–10.

(142) Lin, C. Y.; Huang, Z.; Wen, W.; Wu, A.; Wang, C.; Niu, L. Enhancing Protein Expression in HEK-293 Cells by Lowering Culture Temperature. *PLoS One* **2015**, *10* (4), e0123562.

(143) Yang, J.-Y.; Zong, C. S.; Xia, W.; Wei, Y.; Ali-Seyed, M.; Li, Z.; Broglio, K.; Berry, D. A.; Hung, M.-C. MDM2 Promotes Cell Motility and Invasiveness by Regulating E-Cadherin Degradation. *Mol. Cell. Biol.* **2006**, *26* (19), 7269–7282.

(144) MERK. MG-132 - CAS 133407-82-6 - Calbiochem https://www.sigmaaldrich.com/catalog/product/mm/474790?lang=pt&region=PT (accessed Aug 13, 2018).

(145) Ladanyi, M.; Lewis, R.; Jhanwar, S. C.; Gerald, W.; Huvos, A. G.; Healey, J. H. MDM2 and CDK4 Gene Amplification in Ewing's Sarcoma. *J. Pathol.* **1995**, *175* (2), 211–217.

(146) Iwakuma, T.; Lozano, G. MDM2, An Introduction. *Mol. Cancer Res.* **2002**, *1*, 993–1000.

(147) Gasco, M.; Shami, S.; Crook, T. The p53 Pathway in Breast Cancer. *Breast Cancer Res.* **2002**, *4* (2), 70–76.

(148) Housman, G.; Byler, S.; Heerboth, S.; Lapinska, K.; Longacre, M.; Snyder, N.; Sarkar, S. Drug Resistance in Cancer: An Overview. *Cancers (Basel).* **2014**, *6* (3), 1769–1792.

(149) World Health Organization. Cancer http://www.who.int/news-room/fact-sheets/detail/cancer (accessed Aug 17, 2018).

(150) Dormoy-Raclet, V.; Markovits, J.; Malato, Y.; Huet, S.; Lagarde, P.; Montaudon, D.; Jacquemin-Sablon, A.; Jacquemin-Sablon, H. Unr, a Cytoplasmic RNA-Binding Protein with Cold-Shock Domains, Is Involved in Control of Apoptosis in ES and HuH7 Cells. *Oncogene* **2007**, *26* (18), 2595–2605.

(151) Backer, M. V.; Backer, J. M.; Chinnaiyan, P. Targeting the Unfolded Protein Response in Cancer Therapy. *Methods Enzymol.* **2011**, *491*, 37–56.

(152) Hamanaka, R. B. PERK and GCN2 Contribute to eIF2 Phosphorylation and Cell Cycle Arrest after Activation of the Unfolded Protein Response Pathway. *Mol. Biol. Cell* **2005**, *16* (12), 5493–5501.